GB04/03231

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed P. Mahoney

Dated 1 July 2004

Patents Form 1/77

Patents Act 1977
(Rule 16)

THE PATENT OFFICE
01 AUG 03
RECEIVED BY FAX

The Patent Office
PATENTS • DESIGNS • COPYRIGHT • TRADE MARKS

04AUG03 E827289-1 D02093
P01/7700 0.00-0318109.6

1/77

Request for grant of a patent

See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form.

The Patent Office

Cardiff Road
Newport
South Wales
NP10 8QQ

1. Your reference — PPD 70298/GB/P — 01 AUG 2003

2. Patent application number *(The Patent Office will fill in this part)* — 0318109.6

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*

SYNGENTA Limited
European Regional Centre
Priestley Road
Surrey Research Park, Guildford,
Surrey, GU2 7YH, United Kingdom

Patents ADP number *(if you know it)*

~~6254007002~~ 8330748001

If the applicant is a corporate body, give the country/state of its incorporation — UNITED KINGDOM

4. Title of the invention — IMPROVEMENTS IN OR RELATING TO ORGANIC COMPOUNDS

5. Name of your agent *(if you have one)*

"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

Kathryn ELLIOTT
Intellectual Property Department
Syngenta Limited
Jealott's Hill International Research Centre
PO Box 3538
Bracknell, Berkshire, RG42 6YA
UNITED KINGDOM

Patents ADP number *(if you know it)* — 8029563001

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application	Date of filing *(day / month / year)*

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:*

a) any applicant named in part 3 is not an inventor, or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.
See note (d))

YES (b)

Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	
Description	31
Claim(s)	02
Abstract	01
Drawing(s)	07

10. If you are also filing any of the following, state how many against each item.

Priority documents	
Translations of priority documents	
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	
Request for preliminary examination and search *(Patents Form 9/77)*	
Request for substantive examination *(Patents Form 10/77)*	
Any other documents *(please specify)*	

11. I/We request the grant of a patent on the basis of this application.

Syngenta Limited

Signature C. Dowling Date 1 Aug 03

Authorised Signatory

12. Name and daytime telephone number of person to contact in the United Kingdom

Clare DOWLING = 01344 414834

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) *If you need help to fill in this form or you have any questions, please contact the Patent Office on 08459 500505.*

b) *Write your answers in capital letters using black ink or you may type them.*

c) *If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.*

d) *If you have answered 'Yes' Patents Form 7/77 will need to be filed.*

e) *Once you have filled in the form you must remember to sign and date it.*

f) *For details of the fee and ways to pay please contact the Patent Office.*

IMPROVEMENTS IN OR RELATING TO ORGANIC COMPOUNDS

The present invention relates to *inter alia*, nematicidal and/or insecticidal proteins, nucleic acid sequences encoding them and transformed plants containing them. In particular the invention relates to transformed crops having increased resistance and/or tolerance to nematodes and/or insects.

Nematodes are essentially simple worms generally having cylindrical bodies that taper towards the head and tail. Nematodes that are of agricultural importance can be subdivided into endoparasites, ectoparasites and ecto-endoparasites of plants. Some are sedentary and others remain mobile as they feed. All use a stylet to pierce plant cell walls and feed by removing plant cell contents before or after plant cell modification. More detail of particular important genera and species, their host ranges and economic importance are defined in well known standard texts. The genera *Heterodera* and *Globodera* cyst nematodes are important crop pests. They include *H. glycines*, (soybean cyst nematode) *H. schachtii* (beet cyst nematode), *H. avenae* (cereal cyst nematode) and potato cyst nematodes *G. rostochiensis* and *G. pallida*. Root-knot nematodes particularly the genus *Meloidogyne*, damage a wide range of crops. Examples are species *M. javanica, M. hapla, M. arenaria* and *M. incognita*. There are many other economically important nematodes. Both the above groups produce swollen sedentary females as do other economic genera including *Rotylenchulus, Nacobbus*, and *Tylenchulus*. Other economic nematodes remain mobile as adult females and many of these cause damage to a wide range of crops. Examples include species of *Ditylenchus, Radopholous, Pratylenchus, Helicotylenchus* and *Hirschmanniella*. Others do not always enter plants but feed from them as ectoparasites. Examples include *Aphelenchoides, Anguina Criconemoides, Criconema Hemicycliophora, Hemicriconemoides, Paratylenchus* and *Belonolaimus*. Among the ectoparasites the genera *Xiphinema, Longidorus, Paralongidorus, Trichodorus* and *Paratrichodrus* have distinctive importance. They cause damage to crops by their feeding but their economic status as crop pests is often due to roles as vectors of *inter alia*, NEPO and TOBRA plant viruses.

Plant parasitic nematodes are reported to be responsible for the world-wide loss of billions of pounds worth of agricultural crops each year and so the professional agriculturist is constantly seeking ways to reduce these losses whilst maintaining a high quality commercial crop.

The present invention therefore seeks to provide, *inter alia*, proteins that are active against nematode and/or insect pests.

According to the present invention there is provided a nematicidal protein comprising or consisting of the sequence of SEQ ID NO: 1.

The present invention further provides a nematicidal protein having at least 70% identity to the protein of SEQ ID NO: 1. In a still further embodiment said protein has at least 75% identity to the protein of SEQ ID NO: 1. In a still further embodiment said protein has at least 80% identity to the protein of SEQ ID NO: 1. In a still further embodiment said protein has at least 85% identity to the protein of SEQ ID NO: 1. In a still further embodiment said protein has at least 90% identity to the protein of SEQ ID NO: 1. In a still further embodiment said protein has at least 91% identity to the protein of SEQ ID NO: 1. In a still further embodiment said protein has at least 92% identity to the protein of SEQ ID NO: 1. In a still further embodiment said protein has at least 93% identity to the protein of SEQ ID NO: 1. In a still further embodiment said protein has at least 94% identity to the protein of SEQ ID NO: 1. In a still further embodiment said protein has at least 95% identity to the protein of SEQ ID NO: 1. In a still further embodiment said protein has at least 96% identity to the protein of SEQ ID NO: 1. In a still further embodiment said protein has at least 97% identity to the protein of SEQ ID NO: 1. In a still further embodiment said protein has at least 98% identity to the protein of SEQ ID NO: 1. In a still further embodiment said protein has at least 99% identity to the protein of SEQ ID NO: 1.

The percentage of sequence identity for proteins is determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the amino acid sequence in the comparison window may comprise additions or deletions (i.e. gaps) as compared to the initial reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical amino acid residue occurs in both sequences to yield the number of match positions, dividing the number of match positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity. When calculating the percentage sequence identity the sequences may be aligned allowing for up to 3 gaps with the *proviso* that in respect of the gaps, a total of not more than 15 amino acid residues is affected. Optimal alignment of sequences for comparison may also be conducted by computerised implementations of known algorithms. In a particular embodiment of the present invention the sequence identity

is calculated using the FASTA version 3 algorithm which uses the method of Pearson and Lipman (Lipman, D.J. and Pearson, W.R. (1985) Rapid and sensitive protein similarity searches and Science. 227:1435-1441 and Pearson, W.R. and Lipman, D.J. (1988) Improved tools for biological sequence comparison. PNAS. 85:2444-2448) to search for similarities between the reference sequence (also termed the query sequence) and any group of sequences (termed further sequences). There are also algorithms available to the person skilled in the art that enable a calculation of the percentage sequence identity between polynucleotide sequences.

The protein may differ from the basic nematicidal protein of SEQ ID NO: 1 by conservative or non-conservative amino acid substitutions. A conservative substitution is to be understood to mean that the amino acid is replaced with an amino acid with broadly similar chemical properties. In particular conservative substitutions may be made between amino acids within the following groups:

(i) Alanine and Glycine;

(ii) Serine and Threonine;

(ii) Glutamic acid and Aspartic acid;

(iii) Arginine and Lysine;

(iv) Asparagine and Glutamine;

(v) Isoleucine and Leucine,

(vi) Valine and Methionine;

(vii) Phenylalanine and Tryptophan.

In general, more conservative than non-conservative substitutions will be possible without destroying the nematicidal properties of the proteins. Suitable variants may be determined by testing nematicidal properties of the peptide using routine methods that are well known to the person skilled in the art.

The present invention still further provides a nematicidal protein as described above having a specific lactose binding ability.

The present invention still further provides a nematicidal protein as described above having the ability to cleave disaccharides and/or polysaccharides. In a further embodiment, the protein is a mannanase.

In one aspect of the present invention, the nematicidal protein described above is obtainable from *Lepista nuda*.

The present invention further provides a polynucleotide which encodes a protein as described above. The person skilled in the art will appreciate that there are a number of polynucleotides which can encode each protein due to the degeneracy of the genetic code. In a particular embodiment of the invention the polynucleotide comprises or consists of the sequence of SEQ ID NO: 2.

The present invention further provides a polynucleotide which is the complement of one which hybridises to a sequence as described above under stringent conditions and wherein said polynucleotide still encodes a protein which is nematicidal. Such stringent conditions are well known to the person skilled in the art and comprise, for example: hybridisation at a temperature of about 65°C in a solution containing 6 x SSC, 0.01% SDS and 0.25% skimmed milk powder, followed by rinsing at the same temperature in a solution containing 0.2 x SSC and 0.1% SDS followed by confirmation that the polynucleotide so identified still encodes an insecticidal and/or nematicidal protein according to the invention. The skilled man may alternatively select the following hybridisation conditions *viz.*, hybridisation at a temperature of between 60°C and 65°C in 0.3 strength citrate buffered saline containing 0.1% SDS followed by rinsing at the same temperature with 0.3 strength citrate buffered saline containing 0.1% SDS followed by confirmation that the polynucleotide so identified, still encodes a nematicidal protein according to the present invention.

Further polynucleotides according to the present invention may be identified from nucleic acid libraries. Suitable oligonucleotide probes may be constructed on the basis of the amino acid sequence information of the proteins according to the present invention and used to screen any such nucleic acid library for the identification of further polynucleotides encoding proteins according to the invention. In a particular embodiment of the present invention the a polynucleotide selected from the group consisiting of SEQ ID NOs: 2, 30, 32, 34, 36 and 37 may be used for the construction of oligonucleotide probes by the skilled person. The person skilled in the art is well versed in methods for the production and screening of nucleic acid libraries and the necessary techniques for the subsequent identification, isolation and sequence determination of polynucleotides which encode further nematicidal proteins in accordance with the present invention. The person skilled in the art will appreciate that alternative methods exist for the identification and characterisation of related sequences from *inter alia*, plant/fungal sources. Such methods include PCR

strategies based on oligonucleotide primers using the sequence information provided herein or from sequences obtainable by the methods described above.

In one aspect of the present invention, the polynucleotide desribed above is obtainable from *Lepista nuda*.

The present invention still further provides a polynucleotide having a first region encoding a protein as described above and a second region encoding a further protein. The regions may be separated by a region which provides for a self-processing polypeptide which is capable of separating the proteins such as the self-processing polypeptide described in US Patent No. 5,846,767 or any similarly functioning element. Alternatively the regions may be separated by a sequence which acts as a target site for an external element which is capable of separating the protein sequences. Alternatively the polynucleotide may provide for a polyprotein which comprises a plurality of protein functions. In a further embodiment of the present invention the proteins of the polyprotein may be arranged in tandem. These polyproteins may comprise the proteins according to the present invention and optionally further proteins such as those encoding any desired agronomic trait, for example, a further agronomic trait selected from the group consisting of herbicide resistance, insect resistance, fungus resistance, nematode resistance, altered stress tolerance, altered yield and altered nutritional content. In a particular embodiment of the present invention the further agronomic trait provides resistance to a herbicide which comprises glyphosate acid or agriculturally acceptable salt thereof. In a particular embodiment the polynucleotide may be used to provide the proteins and proteins compositions of the present invention that are described herein.

The present invention still further provides a method of evolving a polynucleotide which encodes a protein having nematicidal properties comprising: (a) providing a population of variants of said polynucleotide and further polynucleotides which encode further proteins, at least one of which is in cell free form; and (b) shuffling said variants and further polynucleotides to form recombinant polynucleotides; and (c) selecting or screening for recombinant polynucleotides which have evolved towards the said nematicidal properties; and (d) repeating steps (b) and (c) with the recombinant polynucleotides according to step (c) until an evolved polynucleotide which encodes a protein having nematicidal properties has been acquired wherein said population of variants in part (a) contains at least one polynucleotide as described in the preceding paragraphs. The methods

for evolving a polynucleotide as described above are well known to the person skilled in the art and are described *inter alia*, in US Patent No. 5,811,238.

The present invention still further provides a construct comprising a polynucleotide as described above. In a further embodiment, the construct comprises a polynucleotide as described above operably linked to a transcription initiation region and a transcription termination region. In a particular embodiment, the transcriptional initiation region and/or the transcription termination region are operable in plants. In a still further embodiment of the present invention the construct may further comprise a region which provides for the targeting of the protein product to a particular location. For example if it is desired to provide the protein outside of the cell then an extracellular targeting sequence may be ligated to the polynucleotide encoding the protein of the present invention. Other examples of targeting include targeting to a specific intracellular organelle or compartment. In a still further embodiment of the present invention the construct may further comprise a region which provides for the retention of the protein at a specific location such as a particular organelle or compartment. For example if it is desired to retain a protein in the endoplasmic reticulum, the KDEL endoplasmic reticulum retention sequence may be used. The present invention still further provides a construct as described above which further comprises a region which provides for the production of a selectable marker. In a particular embodiment, the selectable marker is provided by the mannose isomerase (Positech™) selection system. The selectable marker may, in particular, confer resistance to kanamycin, hygromycin or gentamycin. Further suitable selectable markers include genes that confer resistance to toxins such as eutypine or that confer resistance to herbicides such as protoporphyrinogen oxidase (PPO) inhibitor-based herbicides or glyphosate-based herbicides. In a particular embodiment, the gene may be a mutated PPO such as the ones described in PCT Publication No. WO 97/32011 and PCT Publication No. WO 95/34659 and found, for example, in Acuron™ technology. Other forms of selection are also available such as hormone based selection systems such as the Multi Auto Transformation (MAT) system of Hiroyrasu Ebinuma *et al.* 1997. PNAS Vol. 94 pp2117-2121; visual selection systems which use the known green fluorescence protein, β glucoronidase and any other selection system such as xylose isomerase and 2-deoxyglucose (2-DOG).

The present invention still further provides a construct as described above wherein the plant operable promoter is selected from the group consisting of: CaMV35S, FMV35S, NOS, OCS, Patatin, E9, alcA/alcR switch, GST switch, RMS switch, oleosin, ribulose

bisphosphate carboxylase-oxygenase small sub-unit, actin 7, Cestrum or root specific promoters including MR7 promoter (maize), Gos 9 (rice), GOS2 promoters, a superMas promoter as described in PCT Publication No. WO 95/14098 and US Patent No. 5,955,646 (for example $(Ocs)_3Mas$), the *Agrobacterium rhizogenes* RolD promoter, nematode feeding site specific promoters or other promoters which are derivable from nematodes or are capable of initiating expression at a location where the nematode feeds. Terminators that can be used in the constructs according to the present invention include Nos, proteinase inhibitor II, CaMV35S and the terminator of a gene of alpha-tubulin (European Patent Application, Publication No. EP 0 652 286 A). It is equally possible to use, in association with the promoter regulation sequence, other regulation sequences which are situated between the promoter and the sequence encoding the protein according to the present invention, such as transcriptional or translational enhancers, for example, tobacco etch virus (TEV) translation activator described in PCT Publication No. WO 87/07644. It is also possible to use, in association with the promoter regulation sequence, other regulation sequences such as introns. The polynucleotide encoding the nematicidal protein according to the invention may also be codon-optimised, or otherwise altered to enhance, for example, transcription once it is incorporated into plant material. Such codon optimisation may also be used to alter the predicted secondary structure of the RNA transcript produced in any transformed cell, or to destroy cryptic RNA instability elements present in the unaltered transcript, thereby increasing the stability and/or availability of the transcript in the transformed cell (Abler and Green. 1996. Plant Molecular Biology (32) pp63-78). Table 1 below shows exemplary codon usage preferences for soyabean.

Table 1. Codon-usage preferences for Soyabean

Amino Acid	Soya Preference
Alanine	GCU
Arginine	AGA
Asparagine	AAC
Aspartic Acid	GAU
Cysteine	UGC
Glutamine	CAA
Glutamic Acid	GAG
Glycine	GGA
Histidine	CAU
Isoleucine	AUU
Leucine	CUU
Lysine	AAG
Methionine	ATG
Phenylalanine	UUC
Proline	CCA
Serine	UCU
Threonine	ACU
Tryptophan	UGG
Tyrosine	UAU
Valine	GUU

The present invention still further provides a host cell comprising a polynucleotide or a construct as described above. Suitable host cells include micro-organisms or cultured cells. Suitable micro-organisms include *Escherichia coli* and *Pseudomonas*. Suitable cultured cells include cultured insect cells, cultured mammalian cells and plant cells. The proteins expressed in the host cells described above may be used in the aspects and embodiments of the invention described herein. In addition, these proteins may be used in the production of antibodies using techniques well known in the art. Antibodies produced using the proteins of the invention are suitable for use in screening assays, for example.

The present invention still further provides a transgenic plant comprising a polynucleotide or a construct as described above.

The present invention still further provides a method of providing a plant or plant part with a nematicidal protein comprising: (a) inserting into the genome of the plant or of plant material a polynucleotide or a construct as described above; and (b) regenerating plants or plant parts therefrom; and (c) selecting those plants or plant parts having said protein. The

said polynucleotide or construct may be incorporated into the cells by plant transformation techniques that are well known to the person skilled in the art. Such techniques include but are not limited to particle mediated biolistic transformation, *Agrobacterium*-mediated transformation, protoplast transformation (optionally in the presence of polyethylene glycols); sonication of plant tissues, cells or protoplasts in a medium comprising the polynucleotide or vector; micro-insertion of the polynucleotide or vector into totipotent plant material (optionally employing the known silicon carbide "whiskers" technique), electroporation and the like.

The present invention still further provides plants or plant parts obtained according to the method of the preceding paragraph. In the context of the present invention, plant parts include, but are not limited to, protoplasts, cells, seeds, fruits, leaves, flowers and the like and any other part of the plant that can be reproduced either sexually, asexually or both. In a particular embodiment the plants or plant parts of the present invention are selected from the group consisting of: melon, mango, soybean, cotton, tobacco, sugar beet, oilseed rape, canola, flax, sunflower, potato, tomato, alfalfa, lettuce, maize, wheat, sorghum, rye, banana, barley, oat, turf grass, forage grass, sugar cane, pea, field bean, rice, pine, poplar, apple, peach, grape, strawberry, carrot, cabbage, onion, citrus, cereal or nut plants or any other horticultural crops. Plants and plant parts in accordance with the present invention show improved resistance or enhanced tolerance to a nematode pest when compared to control-like or wild-type plants. Resistance may vary from a slight increase in tolerance to the pest to total resistance so that the plant is unaffected by the presence of pest (where the pest is severely inhibited or killed).

The invention still further provides plants or plant parts as described above which comprise a further agronomic trait selected from the group consisting of herbicide resistance, insect resistance, fungus resistance, nematode resistance, altered stress tolerance, altered yield, altered nutritional content and any other desired agronomic trait. In a particular embodiment of the present invention the further agronomic trait provides resistance to a herbicide which comprises glyphosate acid or agriculturally acceptable salt thereof.

The invention still further provides plants or plant parts as described above which comprise a pharmaceutical trait.

The present invention still further provides the use of a polynucleotide or a construct as described above in a method of producing plants which are resistant and/or tolerant to nematodes.

The present invention still further provides the use of at least one protein as described above as an active ingredient in the production of a nematicide. In a still further embodiment of the present invention the nematicide further comprises an agriculturally acceptable carrier and/or a diluent and/or a nemtatode attractant and may be formulated for use as a spray.

The present invention still further provides a method of controlling nematodes comprising providing at a locus where said nematodes feed at least one protein as described above.

The present invention still further provides a method of controlling nematodes comprising providing at a locus where said nematodes feed a mannanase.

The present invention still further provides a composition comprising a nematicidally effective amount of at least one protein as described above and an agriculturally acceptable carrier and/or a diluent and/or a nematode attractant. In a particular embodiment, the composition comprises a combination of at least two of the proteins described above.

The present invention still further provides a composition comprising a nematicidally effective amount of at least one protein as described above and at least one further protein and an agriculturally acceptable carrier and/or a diluent and/or a nematode attractant. In a particular embodiment, the further protein may be a herbicidal protein, a fungicidal protein, an insectidal protein or another nematicidal protein. The nematicidal proteins according to the present invention may also be combined in application with other agrochemicals such as herbicides, fungicides, insecticides and nematicides. Examples of possible mixture partners include insecticidal and nematicidal lectins, insecticidal protease inhibitors and insecticidal proteins derived from species of the *Bacillus thurigiensis*, *Xenorhabdus nematophilus*, or *Photorabdus luminescens* and other chemicals for example pyrethroids, carbamates, imidacloprid, organochlorines, macromolecules such as spinosad abamectin or emamectin.

The present invention still further provides a recombinant micro-organism which provides for production of a protein as described above. In a further embodiment of the invention the recombinant micro-organism is an endophyte or a *Pseudomonas sp.* An endophyte is generally accepted within the art as a micro-organism having the ability to enter into non-pathogenic endosymbiotic relationships with a plant host. A method of endophyte-enhanced protection of plants has been described in a series of patent applications by Crop Genetics International Corporation (for example, PCT Publication No. WO 90/13224, European Patent Publication No. EP 0 125 468 B1, PCT Publication No. WO 91/10363, PCT Publication No. WO 87/03303). PCT Publication No. WO 94/16076

describes the use of endophytes which have been genetically modified to express a plant-derived insecticidal peptide.

In a further aspect of the present invention, the proteins described above are insecticidal proteins.

The present invention still further provides a recombinant baculovirus which provides for production of a protein as described above.

In a further aspect of the present invention there is provided the use of a recombinant micro-organism or a baculovirus as described above in a method of controlling insects.

The present invention also envisages that the proteins of the invention are active against nematodes which infect animals or humans. Accordingly, there is provided a pharmaceutical composition comprising at least one protein as described above and a pharmaceutically acceptable carrier.

In a further aspect of the present invention, there is provided a protein as described above for use as a pharmaceutical or a vaccine.

In a still further aspect the present invention provides for the use of a protein as described above in the manufacture of a medicament for vaccination against or treatment of animal or human nematode infection.

The nematodes to be controlled by the proteins of the present invention may include but are not limited to: *Heterodera sp.*; *H. schachtii*; *H. glycines*; *H. avenae*; *Meliodogyne sp.*; *M. incognita*; *M. javanica*; *M. hapla*; *M. arenaria*; *Globodera sp.* including *G. rostochiensis*; *G. pallida*; *Tylenchulus sp.*; *Rotylenchulus sp.*; *Xiphinema sp.*; *Longidorus sp.*; *Trichodorus sp.*; *Paratrichodorus sp.*; *Scutellonema sp.*; *Helicotylenchus sp.*; *Hirschmanniella sp.*; *Pratylenchus sp.*; *Ditylenchus sp.*; *Radolpholus sp.*; *Aphelenchoides*; *Anguina Criconemoides*; *Criconema Hemicycliophora*; *Hemicriconemoides* and *Belonolaimus.*

The insects to be controlled by the proteins of the present invention may include but are not limited to the plant chewing insects and the plant chewing stages of insects including: *Coleoptera*, *Lepidoptera*, *Orthoptera* and *Drosophila*, including, but not limited to: *Acanthoscelides obtectus*, *Bruchus sps.*, *Callosobruchus sps.* (bruchid beetles), *Agriotes sps.* (wireworms), *Amphimallon sps.* (chafer beetles), *Anthonomus grandis* (cotton boll weevil), *Ceutorhynchus assimilis* (cabbage seed weevil), *Cylas sps.* (sweet potato weevils), *Diabrotica sps.* (corn root worms), *Epicauta sps.* (black blister beetles), *Epilachna sps.* (melon beetles etc.), *Leptinotarsa decemlineata* (Colorado potato beetle) *Meligisthes sps.*

(blossom beetles), *Melolontha sps.* (cockchafers), *Phyleotreta sps.*, *Psylliodes sps.* (flea beetles), *Popillia japonica* (Japanese beetle), *Scolytus sps.* (bark beetles), *Sitophilus sps.* (grain weevils), *Tenebrio molitor* (yellow mealworm), *Tribolium sps.* (flour beetles), *Trogoderma granarium* (Khapra beetle), *Acleris sps.* (fruit tree tortrixs), *Acraea acerata* (sweet potato butterfly), *Agrotis sps.* (cutworms), *Autographa gamma* (silver-Y moth), *Chilo sps.* (stalk borers), *Cydia pomonella* (codling moth), *Diparopsis sps.* (red bollworms), *Ephestia sps.* (warehouse moths), *Heliothis sps.*, *Helicoverpa sps.* (budworms, bollworms), *Mamestra brassicae* (cabbage moth), *Manduca sps.* (hornworms), *Maruca testulalis* (mung moth), *Mythimna sps.* (cereal armyworms), *Ostrinia nubilalis* (European corn borer), *Pectinophora gossypiella* (pink bollworm), *Phthorimaea operculella* (potato tuber moth), *Pieris brassicae* (large white butterfly), *Pieris rapae* (small white butterfly), *Plodia interpunctella* (Indian grain moth), *Plutella xylostella* (diamond-back moth), *Sitatroga cerealella* (Angoumois grain moth), *Spodoptera sps.* (armyworms), *Trichoplusia ni* (cabbage semilooper), *Acheta sps.* (field crickets), *Gryllotalph sps.* (mole crickets), *Locusta migratoria* (migratory locust), *Schistocerca gregaria* (desert locust), *Acrythosiphon pisum* and *Drosophila melanogaster.*

The invention will now be described by way of the following non-limiting examples in combination with the following Figures and Sequence Listing of which:

Figure 1: Sequence of clone pLnuda-c18 (SEQ ID NO: 30) aligned with the corresponding peptide sequence (SEQ ID NO: 31). The nucleotide is shown as a double stranded polynucleotide sequence (SEQ ID NO: 30 and its complement, SEQ ID NO: 38) with the corresponding translated peptide sequence (SEQ ID NO: 31) aligned above the coding region. The positions of the *L. nuda* tryptic peptide sequences (SEQ ID NOs: 3, 4, 5 and 6) and the 5' UTR are indicated.

Figure 2: Consensus sequence of six 5' RACE cDNA clones (Lnuda-c25, c26, c27, c30, c31 and c32; SEQ ID NO: 32) aligned with the corresponding peptide sequence (SEQ ID NO: 33). The nucleotide is shown as a double stranded polynucleotide sequence (SEQ ID NO: 32 and its complement, SEQ ID NO: 39) with the corresponding translated peptide sequence (SEQ ID NO: 33) aligned above the coding region. The positions of the translation start codon and the 5' UTR are indicated.

Figure 3: Consensus sequence of seven 3' RACE cDNA clones (Lnuda-c37, c38, c40, c41, c42, c47 and c48; SEQ ID NO: 34) aligned with the corresponding peptide sequence (SEQ ID NO: 35). The nucleotide is shown as a double stranded polynucleotide sequence (SEQ ID NO: 34 and its complement, SEQ ID NO: 40) with the corresponding translated peptide sequence aligned above the coding region. The positions of three polymorphisms, the translation stop codon, the 3' UTR and the poly A tail are indicated.

Figure 4: Sequence of full-length cDNA clone Lnuda-c57 (SEQ ID NO: 36) aligned with the corresponding protein sequence (SEQ ID NO: 1). The nucleotide is shown as a double stranded polynucleotide sequence (SEQ ID NO: 36 and its complement, SEQ ID NO: 41) with the corresponding translated protein sequence aligned above the coding region. The positions of three polymorphisms, the translation start codon and the translation stop codon are indicated.

Figure 5: Schematic representation of the positions of the translation start and stop codons and the four introns and four exons of the on a full-length gDNA map of the *L.nuda* gene for the nematicidal protein of SEQ ID NO: 1. In this figure, the translation start and stop codons are indicated by vertical black bars, the exons by striped arrows, the introns by a transparent bar and flanking regions (the 5' and 3' UTRs) by a horizontal black bar.

Figure 6: Consensus sequence of twenty full-length gDNA clones (SEQ ID NO: 37) aligned with the correseponding protein sequence (SEQ ID NO: 1). The nucleotide is shown as a double stranded polynucleotide sequence (SEQ ID NO: 37 and its complement, SEQ ID NO: 42) with the corresponding translated protein sequence aligned above the coding region. The positions of the translation start and stop codons and the introns and exons are indicated.

SEQ ID NO: 1 is an exemplary nematicidal protein of the invention.
SEQ ID NO: 2 is a polynucleotide encoding the protein of SEQ ID NO: 1.
SEQ ID NO: 3 to SEQ ID NO: 11 are protein fragments identified by de novo protein sequencing.
SEQ ID NO: 12 to SEQ ID NO: 29 are primers used in the present invention.
SEQ ID NO: 30 is the sequence of clone pLnuda-c18.

SEQ ID NO: 31 is the peptide sequence corresponding to SEQ ID NO: 30.

SEQ ID NO: 32 is the sequence identified in six 5' RACE cDNA clones (Lnuda-c25, c26, c27, c30, c31 and c32).

SEQ ID NO: 33 is the peptide sequence corresponding to SEQ ID NO: 32.

SEQ ID NO: 34 is the sequence identified in seven 3' RACE cDNA clones (Lnuda-c37, c38, c40, c41, c42, c47 and c48).

SEQ ID NO: 35 is the peptide sequence corresponding to SEQ ID NO: 34.

SEQ ID NO: 36 is the sequence of full-length cDNA clone Lnuda-c57.

SEQ ID NO: 37 is the consensus sequence of 20 full-length gDNA clones.

SEQ ID NO: 38 to SEQ ID NO: 42 are complementary sequences to SEQ ID NO: 30, 32, 34, 36 and 37, respectively.

EXAMPLES

General molecular biology methods are carried out according to Sambrook *et al.* (1989) 'Molecular cloning: A laboratory Manual, 2nd Edition. Cold Spring Harbour Lab. Press.

Example 1: Protein sequencing

a) De novo protein sequencing of *L. nuda* (15kDa band)

A crude extract from *L. nuda* was shown to have nematicidal activity using the method of Example 6. A 15kDa protein band was purified from this extract using the method of Example 5. Tryptic peptides from this protein sample were prepared for nanaospray MS/MS sequencing by digestion using trypsin of either solution or polyacrylamide gel slices containing the protein sample. After digestion with trypsin the peptides were desalted using C_{18} zip tips ™. The peptides were eluted from the zip tips using 50:50:0.1 acetonitrile: water: formic acid and directly transferred into a nano electrospray needle. Mass spectrometry of the desalted peptides was carried out using a micromass Q-TOF 2 mass spectrometer. The mass spectrometer was calibrated using the product ions from the fragmentation of the doubly charged ion of glu-fibrinopeptide B in 50:50:0.1 acetonitrile: water: formic acid at 500fmol/µl. Ions from the nanospray needle were detected in the Q-TOF mass spectrometer running in survey mode. Multiply charged ions were then selected one at a time and fragmented by collision with argon gas to form

fragment ions. The fragment ion spectra obtained are referred to as MS/MS spectra. The most appropriate collision energy for each peptide was selected manually by the operator. Under the low energy conditions used the predominant cleavage occurs at the peptide bond generating 2 types of ions, the y-ions formed from the C-terminus of the peptide and the b-ions generated from the N-terminus of the peptide. By measuring the difference in masses between different ions in the MS/MS spectra it is possible to propose amino acid sequences. Using the above approach the following peptide sequences were proposed:

E(Q/K)ELV(Q/K)SG(Q/K)TYLLTNAK (SEQ ID NO: 3)

LVAVTTPVEWHLWHDEVDHT(366.21) (SEQ ID NO: 4)

WSSEMYLGLNGSPSDGTK (SEQ ID NO: 5)

AVTTPVEWHLWHDEVD(623.23) (SEQ ID NO: 6)

SGNLGLYF or FYLGLNGS (SEQ ID NO: 7 and SEQ ID NO: 8, respectively)

(243.19)TVDLS(1304.65) (SEQ ID NO: 9)

(1716.75)(Q/K)SAAPGSSHTTGEYTWK (SEQ ID NO: 10)

(3821.49)NSVYTWK (SEQ ID NO: 11)

Non standard nomenclature used is as follows:

L	means Leucine or Isoleucine
(Q/K)	means Glutamine or a Lysine at this position
(623.23)	means unable to assign remaining 623.23 dalton fragment
F	means Phenylalanine or oxidized Methionine at this position

b) Edman sequencing of *L. nuda* (15kDa band)

Whole protein samples that had been blotted onto PVDF were N-terminally sequenced by Edman degradation using an ABI Procise cLC Sequencer using programmes supplied by the manufacturer.

The result indicated that the protein is blocked. By comparison of the MS/MS sequencing data with the predicted protein sequence there was a clear indication that the protein starts with an acetylated Serine i.e the first Methionine residue is missing.

Example 2: Cloning of cDNA and gDNA

a. Cloning of Partial cDNA

Nested degenerate oligonucleotides were designed to proposed tryptic peptide sequences generated from MS/MS derived data as detailed in Example 1. The following tryptic peptide sequences were used:

Peptide 1 (SEQ ID NO: 3)	E(Q/K)ELV(Q/K)SG(Q/K)TYLLTNAK	m.wt. = 1920.86
Peptide 2 (SEQ ID NO: 4)	LVAVTTPVEWHLWHDEVDHT(366.21)	m.wt. = 2749.36
Peptide 3 (SEQ ID NO: 5)	WSSEMYLGLNGSPSDGTK	m.wt. = 1927.76
Peptide 4 (SEQ ID NO: 6)	AVTTPVEWHLWHDEVD(623.23)	m.wt. = 2537.13

The following nested oligonucleotides were designed:

Lnuda-1F	GGIMARACITAYHTIHTIAC (SEQ ID NO: 12)
Lnuda-2F	ARGARHTIGTIMARWSIGG (SEQ ID NO: 13)
Lnuda-1R	GCRTTIGTIADIADRTAIGT (SEQ ID NO: 14)
Lnuda-2R	TKICCISWYTKIACIADYTC (SEQ ID NO: 15)
Lnuda-3F	GTIGARTGGCAYHTITGGCA (SEQ ID NO: 16)
Lnuda-4F	ACIACICCIGTIGARTGGCA (SEQ ID NO: 17)
Lnuda-3R	TGRTCIACYTCRTCRTGCCA (SEQ ID NO: 18)
Lnuda-4R	TCRTGCCAIADRTGCCAYTC (SEQ ID NO: 19)

A 418 bp PCR product was amplified from an amplified *L. nuda* cultivated mushroom cDNA library in λZapII (Stratagene #200400/200401/200450) using nested degenerate Lnuda primers in combination with nested T3 promoter specific primers, as follows:

Primary PCR:	Lnuda-3R + T3 (AATTAACCCTCACTAAAGGG; SEQ ID NO: 20)
Secondary PCR:	Lnuda-4R + T3-nest (ACTAAAGGGAACAAAAGCTGG; SEQ ID NO: 21)

25 µl PCR amplification reactions were run using puRe Taq™Ready-To-Go™PCR Beads (Amersham Biosciences #27-9558-01) incorporating 300ng of degenerate primer,

100ng of specific primer and 1µl of cDNA library template under the following conditions: 95°C for 5 minutes, 35 cycles of 95°C for 30 seconds, 50°C for 30 seconds, 72°C for 90 seconds, then 72°C for 6 minutes (Techne Progene thermal cycler). 1µl of primary PCR was used as template for secondary PCR.

The 418 bp PCR amplicon was cloned directly into pCR® 2.1 TOPO® using TOPO TA Cloning Kit (Invitrogen 45-0641) and sequenced by automated cycle sequencing from M13 forward (GTAAAACGACGGCCAG; SEQ ID NO: 22) and reverse (CAGGASAACAGCTATGAC; SEQ ID NO: 23) primers.

The sequence of the 418 bp *L.nuda* PCR amplicon (clone pLnuda-c18) was determined to be:

TACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCA
GGAATTCGGCACGAGGAACTTTCTGCCTCGTTTTTTTGCTCCTACTGTTTTTCTCTTCCAGTTTCTACCATGTC
GCAAGAAATTGTTCAATCAGGACAAACCTACATCATCACTAACGCCAAATCCGGCACAGTTGTTGACCTTTCGG
GCGAAGACAACAAATCTATTATTGGATTTCCCAAGCATGGAGGAACAAATCAGAGGTGGACCCTCAACTGGACA
GGGAAGAGTTGGACTTTCCGCTCCGTTTCTTCTGAAATGTATCTTGGCCTGAATGGCTCGCCGTCTGATGGAAC
AAAACTGGTAGCCGTGACCACCCCTGTTGAGTGGCGCATCTGGCACGA (SEQ ID NO: 30)

This nucleotide sequence, when translated from the open reading frame starting at the ATG codon at nucleotides 144 to 146, corresponds to the following peptide sequence:

MSQEIVQSGQTYIITNAKSGTVVDLSGEDNKSIIGFPKHGGTNQRWTLNWTGKSWTFRSVSSEMYLGLNGSPSD
GTKLVAVTTPVEWRIWH (SEQ ID NO: 31)

Figure 1 shows this nucleotide sequence (pLnuda-c18; SEQ ID NO: 30) aligned with the corresponding peptide sequence (SEQ ID NO: 31). The nucleotide is shown as a double stranded polynucleotide sequence (SEQ ID NO: 30 and its complement, SEQ ID NO: 38) with the corresponding translated peptide sequence (SEQ ID NO: 31) aligned above the coding region. The positions of the *L. nuda* tryptic peptide sequences (SEQ ID NOs: 3, 4, 5 and 6) and the 5' UTR are indicated. All major tryptic peptides are represented on the 418bp cDNA clone thereby confirming authenticity.

b. Cloning of *L. nuda* 5' & 3' cDNA ends

L. nuda gene specific nested forward and reverse primers were designed in order to PCR amplify both 5' and 3' cDNA ends:

Lnuda-S1	TCATCACTAACGCCAAATCCG (SEQ ID NO: 24)
Lnuda-S2	TTGTTGACCTTTCGGGCGAAG (SEQ ID NO: 25)
Lnuda-S3	TTCAGAAGAAACGGAGCGG (SEQ ID NO: 26)
Lnuda-S4	TCCAACTCTTCCCTGTCCAG (SEQ ID NO: 27)

Using *L. nuda* total RNA, 5' and 3' cDNA ends were PCR amplified using First Choice® RLM-RACE Kit (Ambion #1700) as per manufacturers instructions.

A 211 bp 5' RACE PCR amplicon and a 493 bp 3' RACE PCR amplicon were cloned directly into pCR® 2.1 TOPO® and sequenced.

L.nuda 5' RACE consensus sequence:

Six 5' RACE cDNA clones were identified (Lnuda-c25, c26, c27, c30, c31 and c32) having the following consensus sequence:

GCCTCGTTTTTTTGCTCCTACTGTTTTTCTCTTCCAGTTTCTACCATGTCGCAAGAAATTGTTCAATCAGGACA
AACCTACATCATCACTAACGCCAAATCCGGCACAGTTGTTGACCTTTCGGGCGAAGACAACAAATCTATTATTG
GATTTCCCAAGCATGGAGGAACAAATCAGAGGTGGACCCTCAACTGGACAGGGAAGAGTTGGA (SEQ ID NO: 32)

This nucleotide sequence corresponds to the following peptide sequence; the ATG translation start codon being found at nucleotides 46 to 48:

MSQEIVQSGQTYIITNAKSGTVVDLSGEDNKSIIGFPKHGGTNQRWTLNWTGKSW (SEQ ID NO: 33)

Figure 2 shows the consensus sequence of the six 5' RACE cDNA clones (Lnuda-c25, c26, c27, c30, c31 and c32; SEQ ID NO: 32) aligned with the corresponding peptide sequence (SEQ ID NO: 33). The nucleotide is shown as a double stranded polynucleotide sequence (SEQ ID NO: 32 and its complement, SEQ ID NO: 39) with the corresponding translated peptide sequence (SEQ ID NO: 33) aligned above the coding region. The positions of the translation start codon and the 5' UTR are indicated.

L.nuda 3' RACE consensus sequence:

Seven 3' RACE cDNA clones were identified (Lnuda-c37, c38, c40, c41, c42, c47 and c48) having the following consensus sequence.

TTGTTGACCTTTCGGGCGAAGACAACAAATCTATTATTGGATTTCCCAAGCATGGAGGAACAAATCAGAGGTGG
ACCCTCAACTGGACAGGGAAGAGTTGGACTTTCCGCTCCGTTTCTTCTGAAATGTATCTTGGCCTGAATGGCTC
GCCGTCTGATGGAACAAAACTGGTAGCCGTGACCACCCCTGTTGAGTGGCACATCTGGCACGACGAAGTTGACC
CTTCAACTTATCGTATCTTTGTACCTTTCACCACATTCAACATGGACCTCTACGCCCARGGTAGTGCCGCCCCT
GGTACGCCTATCACAACTTGGTATACATGGAAGGGYATCCACCAAACGTGGAGGTTTGAACTAGCTTAGGKTCA
GGTTTCGGATGTAATTTGTGTGTGTAAATCTTCTTGGACCATGTTGTGCTTTTATTGTACTCCGCTTGTTATCA
TTATACCCACCTATGTTGCAACATCTTTTTGGATCCCAAAAAAAAAAAA (SEQ ID NO: 34)

This nucleotide sequence corresponds to the following peptide sequence, translation being carried out starting from the GTT codon at positions 3 to 5 and finishing at the TAG codon at positions 362 to 364:

VDLSGEDNKSIIGFPKHGGTNQRWTLNWTGKSWTFRSVSSEMYLGLNGSPSDGTKLVAVTTPVEWHIWHDEVDP
STYRIFVPFTTFNMDLYAQGSAAPGTPITTWYTWKGIHQTWRFELA (SEQ ID NO: 35)

Two polymorphisms were detected within the coding region (A/G and C/T; shown as R and Y in SEQ ID NO: 34) and one outside of the coding region (T/G; shown as K in SEQ ID NO: 34). The two polymorphisms within the coding region are in the 3rd position and therefore the amino acid sequence is not affected.

Figure 3 shows the consensus sequence of the seven 3' RACE cDNA clones (Lnuda-c37, c38, c40, c41, c42, c47 and c48; SEQ ID NO: 34) aligned with the corresponding peptide sequence (SEQ ID NO: 35). The nucleotide is shown as a double stranded polynucleotide sequence (SEQ ID NO: 34 and its complement, SEQ ID NO: 40) with the corresponding translated peptide sequence aligned above the coding region. The positions of three polymorphisms, the translation stop codon, the 3' UTR and the poly A tail are indicated.

c. Cloning of full-length *L. nuda* cDNA

L. nuda gene specific primers were designed to conserved 5' and 3' UTR regions in order to PCR amplify full-length clones using cDNA template as generated by the RLM-RACE Kit:

Lnuda-S5 TCTCTTCCAGTTTCTACCATG (SEQ ID NO: 28)

Lnuda-S6 ACAAATTACATCCGAAACCTG (SEQ ID NO: 29)

25 µl PCR amplification reactions were run using puRe Taq™Ready-To-Go™PCR Beads (Amersham Biosciences #27-9558-01) incorporating 100ng of each specific primer and 1µl of cDNA template under the following conditions: 95°C for 5 minutes, 35 cycles of 95°C for 30 seconds, 62°C for 30 seconds, 72°C for 90 seconds, then 72°C for 6 minutes (Techne Progene thermal cycler).

A 491 bp PCR amplicon was cloned directly into pCR® 2.1 TOPO® using TOPO TA Cloning Kit (Invitrogen 45-0641) and sequenced by automated cycle sequencing from M13 forward (GTAAAACGACGGCCAG; SEQ ID NO: 22) and reverse (CAGGAAACAGCTATGAC; SEQ ID NO: 23) primers.

A total of 17 full-length *L. nuda* cDNA clones were sequenced. The sequence of clone Lnuda-c57, which is representative of the consensus sequence, is given below:

TCTCTTCCAGTTTCTACCATGTCGCAAGAAATTGTTCAATCAGGACAAACCTACATCATCACTAACGCCAAATC
CGGCACAGTTGTTGACCTTTCGGGCGAAGACAACAAATCTATTATTGGATTTCCCAAGCATGGAGGAACAAATC
AGAGGTGGACCCTCAACTGGACAGGGAAGAGTTGGACTTTCCGCTCCGTTTCTTCTGAAATGTATCTTGGCCTG
AATGGCTCGCCGTCTGATGGAACAAAACTGGTAGCCGTGACCACCCCTGTTGAGTGGCACATCTGGCACGACGA
AGTTGACCCTTCAACTTATCGTATCTTTGTACCTTTCACCACATTCAACATGGACCTCTACGCCCAAGGTAGTG
CCGCCCCTGGTACGCCTATCACAACTTGGTATACATGGAAGGGCATCCACCAAACGTGGAGGTTTGAACTAGCT
TAGGGTCAGGTTTCGGATGTAATTTGT (SEQ ID NO: 36)

The open reading frame within this region coding for a protein product is given below:

ATGTCGCAAGAAATTGTTCAATCAGGACAAACCTACATCATCACTAACGCCAAATCCGGCACAGTTGTTGACCT
TTCGGGCGAAGACAACAAATCTATTATTGGATTTCCCAAGCATGGAGGAACAAATCAGAGGTGGACCCTCAACT
GGACAGGGAAGAGTTGGACTTTCCGCTCCGTTTCTTCTGAAATGTATCTTGGCCTGAATGGCTCGCCGTCTGAT
GGAACAAAACTGGTAGCCGTGACCACCCCTGTTGAGTGGCACATCTGGCACGACGAAGTTGACCCTTCAACTTA

TCGTATCTTTGTACCTTTCACCACATTCAACATGGACCTCTACGCCCAAGGTAGTGCCGCCCCTGGTACGCCTA
TCACAACTTGGTATACATGGAAGGGCATCCACCAAACGTGGAGGTTTGAACTAGCTTAG (SEQ ID NO: 2)

The *L. nuda* gene is comprised of a 429 bp open reading frame (SEQ ID NO: 2), encoding 142 amino acids (SEQ ID NO: 1), thereby giving a predicted molecular weight of 15962.85 Daltons. There appears to be only one gene encoding for the *L. nuda* nematicidal protein of the present invention.

The confirmed protein sequence coded for by SEQ ID NO: 2, and nucleotides 19 to 447 of SEQ ID NO: 35, is as follows:

MSQEIVQSGQTYIITNAKSGTVVDLSGEDNKSIIGFPKHGGTNQRWTLNWTGKSWTFRSVSSEMYLGLNGSPSD
GTKLVAVTTPVEWHIWHDEVDPSTYRIFVPFTTFNMDLYAQGSAAPGTPITTWYTWKGIHQTWRFELA (SEQ ID NO: 1)

This protein is a non-secretory protein as predicted by SignalP. When BlastP searched, there was only one significant match: 55% identity to an *Agaricus bisporus* putative mannanase. Over the full protein length, there is 52.1% identity to the *A.bisporus* putative mannanase.

Figure 4 shows the sequence of the full-length cDNA clone Lnuda-c57 (SEQ ID NO: 36) aligned with the corresponding protein sequence (SEQ ID NO: 1). The nucleotide is shown as a double stranded polynucleotide sequence (SEQ ID NO: 36 and its complement, SEQ ID NO: 41) with the corresponding translated protein sequence aligned above the coding region. The positions of three polymorphisms, the translation start codon and the translation stop codon are indicated.

d. Cloning of full-length L.nuda gDNA

Primers Lnuda-S5 and Lnuda-S6 were used to PCR amplify the full-length coding region from *L.nuda* cultivated mushroom genomic DNA (gDNA). Amplification conditions were as described above.

A resulting 706 bp amplicon was TOPO TA cloned and sequenced, as previously described. The consensus sequence for 20 positive clones is shown below:

TCTCTTCCAGTTTCTACCATGTCGCAAGAAATTGTTCAATCAGGACAAACCTACATCATCACTAACGCCAAATC
CGGCACAGTTGTTGACCTTTCGGGCGAAGACAACAAATCTAGTAAGTCGTTTTTAGTCCCATGTTTTTTTTTGT
CAAAAAAAATTGACTGACATATTTTGTCTCCAGTTATTGGATTTCCCAAGCATGGAGGAACAAATCAGAGGGTA
GGTCTAGAAATGCACCTCGTTAATATTGGTTTTTATTGACATTCATGAACAGTGGACCCTCAACTGGACAGGGA
AGAGTTGGACTTTCCGCTCCGTTTCTTCTGAAATGTATCTTGGCCTGAATGGCTCGCCGTCTGATGGAACAAAA
CTGGTAGCCGTGACCACCCCTGTTGAGTGGCACATCTGGCACGACGAAGTTGACCCTTCAACTTATCGGTGAGT
CCCCTAAATATTACTTGCTTGTGGTTCATACTAATACGTCGTTCGAAGTATCTTTGTACCTTTCACCACATTCA
ACATGGACCTCTACGCCCAGGGTAGTGCCGCCCCTGGTACGCCTATCACAACTTGGTATACATGGAAGGGTATC
CACCAAACGTGGAGGTTTGAACTAGGTAGGGCTTGCGATCTCACCCGGATCCTCCATGAACTAATGTGATCACG
TCGTGTTCTAGCTTAGGTTCAGGTTTCGGATGTAATTTGT (SEQ ID NO: 37)

No polymorphisms were detected. The amino acid sequence is 100% identical to that translated from the cDNA, as predicted. Introns were mapped according to the GT:AG rule. Four intronic regions could be mapped to the gDNA sequence:

Intron 1	66 bp
Intron 2	55 bp
Intron 3	54 bp
Intron 4	60 bp

Intron 4 splices within the very last codon (A) immediately upstream of the TAG stop codon. Figure 5 shows a schematic representation of the positions of the translation start and stop codons and the four introns and four exons on a full-length gDNA map of the *L.nuda* gene encoding for the nematicidal protein of SEQ ID NO: 1. Figure 6 shows the consensus sequence of the twenty full-length gDNA clones (SEQ ID NO: 37) aligned with the correseponding protein sequence (SEQ ID NO: 1). The nucleotide is shown as a double stranded polynucleotide sequence (SEQ ID NO: 37 and its complement, SEQ ID NO: 42) with the corresponding translated protein sequence aligned above the coding region. The positions of the translation start and stop codons and the introns and exons are indicated.

Example 3: Construction of vectors.

The person skilled in the art is capable of producing the polynucleotides and polynucleotide sequences according to the invention utilising standard molecular biology cloning techniques. In addition, various polynucleotides based on the sequence information

as listed in the Sequence Listing and described above can be synthesised chemically using standard techniques well known to the person skilled in the art.

Basic vectors are available or may be constructed which are suitable for assembling the polynucleotides and polynucleotide sequences according to the invention. The component parts of the polynucleotides may be synthesised or may be cloned from other vectors or libraries containing said sequences. The person skilled in the art may generate probes based on the information presented in the Sequence Listing to isolate the sequences according to the invention also using standard techniques. The vectors may be used for DNA work (sequencing, mutagenesis), for *in vitro* production and for plant transformation, following methods well known in the art. When providing vectors for plant transformation which utilise *Agrobacterium*, it is preferred that the sequences according to the invention are inserted between the border regions of a single T-DNA region. *Agrobacterium* may be transformed in accordance with methods which are well known to the person skilled in the art.

Example 4: Transformation

The vectors are transformed into plant cells using standard procedures. Any transformation method suitable for the target plant or plant cells may be employed, including, but not limited to, *Agrobacterium*-mediated transformation (for example, infection by *Agrobacterium tumefaciens* containing recombinant Ti plasmids, infection with *Agrobacterium rhizogenes* transformed with binary vectors to produce transgenic hairy root cultures (Cho H-J *et al.* Planta 210, 195-204)), particle mediated biolistic transformation, protoplast transformation (optionally in the presence of polyethylene glycols), electroporation, sonication of plant tissues, cells or protoplasts in a medium comprising the polynucleotide or vector, microinjection of plant cells and protoplasts, bacterial bombardment, micro-insertion of the polynucleotide or vector into totipotent plant material (optionally employing the known silicon carbide "fibre" or "whisker" method) and pollen tube transformation. The transformed cells may then in suitable cases be regenerated into whole plants in which the new nuclear material is stably incorporated into the genome. Both transformed monocot and dicot plants may be obtained in this way. Full details of the methods of transformation are known to the person skilled in the art.

Example 5: Protein purification

Approximately 100 gram quantities of *L. nuda* were frozen in liquid nitrogen and ground using a pestle and mortar. Ground mushroom was placed into a flat-bottomed dish with plastic, perforated lid and was then freeze dried for two days until all moisture had been removed from the sample.

The lyophilised mushroom was removed from the container and weighed. 20mls of extraction buffer was added for every 1g of lyophilised mushroom. Protease inhibitor cocktail containing a mixture of water-soluble protease inhibitors with broad specificity for the inhibition of serine, cysteine, aspartic, and metallo-proteases (containing 4-(2-aminoethyl) benzenesulfonyl fluoride (AEBSF), E-64, bestatin, leupeptin, aprotinin, and sodium EDTA) was added (1 pot of protease inhibitor cocktail was made up to 100ml with the addition of ultra pure water). 1ml of protease inhibitor cocktail was used for every 100ml of mushroom extract. The extract was incubated overnight at 4°C with stirring to ensure the full extraction of all proteins.

The particulates were then removed by centrifugation. Extract was centrifuged at as high a speed as the rotor allowed, and the supernatant removed. The supernatant was then filtered to ensure all particulates were removed.

In order to concentrate and remove other contaminants, ammonium sulphate was added to the extract to 90% saturation. This was incubated at 4°C with rolling for 1 hour. The protein was then spun down using ss34 rotor at 20,000 rpm for 20mins. The supernatant was removed and the precipitated protein was re-solubilised in 20mM Tris (pH 8.0) in order to give to solution some buffering capacity.

The resuspended extract was dialysed into Lactosyl Agarose Buffer A (PBS pH 7.5) using 3kDa dialysis membrane. This ensured removal of the remaining ammonium sulphate and allowed further purification using lactosyl agarose. This purification was carried out as follows. 2mls of 50% lactosyl-agarose slurry (lactosyl agarose from Sigma in 0.15M NaCl and 0.01M sodium phosphate, pH 6.8) was pipetted into 15ml Falcon tube. The beads were washed 4 times with 10mls of PBS. After each wash, the beads were pelleted by spinning at 3500 rpm for 4mins (speed is important as higher speeds may disrupt the integrity of the beads) and as much supernatant as possible was removed without disturbing the beads. 10mls of mushroom extract was added to the tube which was then incubated overnight to allow lectin binding. The beads were then pelleted again and the supernatant is removed as Unbound fraction. The beads were then washed and pelleted again three times with 10mls of PBS solution to remove the remaining unbound proteins. 10mls of 0.4M α-lactose was then

added and the tube was incubated at 4°C with rolling for 1 hour. The beads were pelleted and the supernatant removed as Bound fraction. 10µl of original extract, unbound and bound was run on Novex 4-12% Bis/tris SDSpage with MES buffer and Mark 12 molecular weight markers.

The lactose bound fraction was concentrated 5 times and loaded onto a superdex 70 (Amersham Biosciences) column pre-equilibrated with PBS. An isocratic gradient was run and 1 mL fractions collected and assayed for activity. This last step purified the nematicidal protein to apparent homogeneity.

Example 6: Nematicidal effect on *Heterodera schachtii* in plants

This screen can be used to assess nematicidal effects of a range of test samples (including crude extracts containing proteins and polypeptides, partially purified proteins or polypeptides, purified proteins or polypeptides and recombinantly expressed proteins or polypeptides). These test samples are delivered directly to feeding sites induced by *Heterodera schachtii* juveniles feeding on *Arabidopsis thaliana* roots.

Sterile *Arabidopsis thaliana* seedlings (2 weeks old) growing in a thin layer of Knop media are inoculated with 100-200 sterile hatched J2s of *Heterodera schachtii* and maintained at 24°C, 16 h light, 8h dark. Between 6-9 d post inoculation individual feeding sites of J3 stage nematodes are injected with approximately 0.2 µl of test sample (concentration of approximately1000 ppm/1 mg/ml) and lucifer yellow dye (10 mM) which enables visualisation of oral uptake by the feeding nematode using an inverted microscope, Nikon Eclipse TE200 with epi-fluorescence microscope attachment. Between 8-20 successful uptakes are carried out for each sample using several individual seedlings. Seedlings with injected feeding sites, from which nematodes have ingested sample, are sandwiched between two glass slides and kept at 24°C, 16 h light, 8h dark for at least 7 d during which time the development of the nematodes are assessed.

The efficacy of the test sample is recorded as the number of male/female/juvenile stage nematodes dead or live after approximately 7-14 d. For all tests effect codes may be attached to the score, including:

Nematode median bulb pumping

Nematode moulting/development

Evidence of vacuolation inside nematode

Example 7: Transformation of soybean (*Glycine max*)

Soybean cotyledon explants are harvested and wounded with a scalpel dipped in a culture of A. *rhizogenes* grown overnight in LB medium containing 150mg/l kanamycin. After wounding, the explants are dried on sterile filter paper and transferred onto co-cultivation media for 3 days. The explants are then transferred onto hormone-free media supplemented with 500 mg/l carbenicillin and 200 mg/l kanamycin. About 10-14 days after root emergence, 1-2cm long root tips are transferred onto fresh new media and regularly sub-cultured every 5-6 weeks. Transgenic hairy roots may be used in feeding assays with nematodes.

Example 8: Transformation of other plants

Other plants, including, but not limited to, *Arabidopsis*, tobacco, sugarbeet, potato and tomato, are transformed using methods well known to those skilled in the art.

Example 9: Analysis of transgenic hairy roots for susceptibility to PPN

Eggs are obtained from females of *H. glycines* collected from soybean plants grown in a greenhouse. Eggs are washed by centrifugation in sterile distilled water, surface sterilized and washed again. About 200 eggs are added to recently sub-cultured transformed root cultures (about 200mg fresh weight). Cultures are incubated at 25°C and about 4-5 weeks after inoculation, scored for presence of female cysts.

Other modifications of the present invention will be apparent to those skilled in the art without departing from the scope of the invention which is defined by the appended claims.

SEQUENCE LISTING

SEQ ID NO: 1

MSQEIVQSGQTYIITNAKSGTVVDLSGEDNKSIIGFPKHGGTNQRWTLNWTGKSWTFRSVSSEMYLGLNGSPSD
GTKLVAVTTPVEWHIWHDEVDPSTYRIFVPFTTFNMDLYAQGSAAPGTPITTWYTWKGIHQTWRFELA

SEQ ID NO: 2

ATGTCGCAAGAAATTGTTCAATCAGGACAAACCTACATCATCACTAACGCCAAATCCGGCACAGTTGTTGACCT
TTCGGGCGAAGACAACAAATCTATTATTGGATTTCCCAAGCATGGAGGAACAAATCAGAGGTGGACCCTCAACT
GGACAGGGAAGAGTTGGACTTTCCGCTCCGTTTCTTCTGAAATGTATCTTGGCCTGAATGGCTCGCCGTCTGAT
GGAACAAAACTGGTAGCCGTGACCACCCCTGTTGAGTGGCACATCTGGCACGACGAAGTTGACCCTTCAACTTA
TCGTATCTTTGTACCTTTCACCACATTCAACATGGACCTCTACGCCCAAGGTAGTGCCGCCCCTGGTACGCCTA
TCACAACTTGGTATACATGGAAGGGCATCCACCAAACGTGGAGGTTTGAACTAGCTTAG

SEQ ID NO: 3

E(Q/K)ELV(Q/K)SG(Q/K)TYLLTNAK

SEQ ID NO: 4

LVAVTTPVEWHLWHDEVDHT

SEQ ID NO: 5

WSSEMYLGLNGSPSDGTK

SEQ ID NO: 6

AVTTPVEWHLWHDEVD

SEQ ID NO: 7

SGNLGLYF

SEQ ID NO: 8

FYLGLNGS

SEQ ID NO: 9

TVDLS

SEQ ID NO: 10

(Q/K)SAAPGSSHTTGEYTWK

SEQ ID NO: 11

NSVYTWK

SEQ ID NO: 12

GGIMARACITAYHTIHTIAC

SEQ ID NO: 13

ARGARHTIGTIMARWSIGG

SEQ ID NO: 14

GCRTTIGTIADIADRTAIGT

SEQ ID NO: 15

TKICCISWYTKIACIADYTC

SEQ ID NO: 16

GTIGARTGGCAYHTITGGCA

SEQ ID NO: 17

ACIACICCIGTIGARTGGCA

SEQ ID NO: 18

TGRTCIACYTCRTCRTGCCA

SEQ ID NO: 19

TCRTGCCAIADRTGCCAYTC

SEQ ID NO: 20

AATTAACCCTCACTAAAGGG

SEQ ID NO: 21

ACTAAAGGGAACAAAAGCTGG

SEQ ID NO: 22

GTAAAACGACGGCCAG

SEQ ID NO: 23

CAGGASAACAGCTATGAC

SEQ ID NO: 24

TCATCACTAACGCCAAATCCG

SEQ ID NO: 25

TTGTTGACCTTTCGGGCGAAG

SEQ ID NO: 26

TTCAGAAGAAACGGAGCGG

SEQ ID NO: 27

TCCAACTCTTCCCTGTCCAG

SEQ ID NO: 28

TCTCTTCCAGTTTCTACCATG

SEQ ID NO: 29

ACAAATTACATCCGAAACCTG

SEQ ID NO: 30

TACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCA
GGAATTCGGCACGAGGAACTTTCTGCCTCGTTTTTTTGCTCCTACTGTTTTTCTCTTCCAGTTTCTACCATGTC
GCAAGAAATTGTTCAATCAGGACAAACCTACATCATCACTAACGCCAAATCCGGCACAGTTGTTGACCTTTCGG
GCGAAGACAACAAATCTATTATTGGATTTCCCAAGCATGGAGGAACAAATCAGAGGTGGACCCTCAACTGGACA
GGGAAGAGTTGGACTTTCCGCTCCGTTTCTTCTGAAATGTATCTTGGCCTGAATGGCTCGCCGTCTGATGGAAC
AAAACTGGTAGCCGTGACCACCCCTGTTGAGTGGCGCATCTGGCACGA

SEQ ID NO: 31

MSQEIVQSGQTYIITNAKSGTVVDLSGEDNKSIIGFPKHGGTNQRWTLNWTGKSWTFRSVSSEMYLGLNGSPSD
GTKLVAVTTPVEWRIWH

SEQ ID NO: 32

GCCTCGTTTTTTTGCTCCTACTGTTTTTCTCTTCCAGTTTCTACCATGTCGCAAGAAATTGTTCAATCAGGACA
AACCTACATCATCACTAACGCCAAATCCGGCACAGTTGTTGACCTTTCGGGCGAAGACAACAAATCTATTATTG
GATTTCCCAAGCATGGAGGAACAAATCAGAGGTGGACCCTCAACTGGACAGGGAAGAGTTGGA

SEQ ID NO: 33

MSQEIVQSGQTYIITNAKSGTVVDLSGEDNKSIIGFPKHGGTNQRWTLNWTGKSW

SEQ ID NO: 34

TTGTTGACCTTTCGGGCGAAGACAACAAATCTATTATTGGATTTCCCAAGCATGGAGGAACAAATCAGAGGTGG
ACCCTCAACTGGACAGGGAAGAGTTGGACTTTCCGCTCCGTTTCTTCTGAAATGTATCTTGGCCTGAATGGCTC
GCCGTCTGATGGAACAAAACTGGTAGCCGTGACCACCCCTGTTGAGTGGCACATCTGGCACGACGAAGTTGACC
CTTCAACTTATCGTATCTTTGTACCTTTCACCACATTCAACATGGACCTCTACGCCCARGGTAGTGCCGCCCCT
GGTACGCCTATCACAACTTGGTATACATGGAAGGGYATCCACCAAACGTGGAGGTTTGAACTAGCTTAGGKTCA
GGTTTCGGATGTAATTTGTGTGTGTAAATCTTCTTGGACCATGTTGTGCTTTTATTGTACTCCGCTTGTTATCA
TTATACCCACCTATGTTGCAACATCTTTTTGGATCCCAAAAAAAAAAAA

SEQ ID NO: 35

VDLSGEDNKSIIGFPKHGGTNQRWTLNWTGKSWTFRSVSSEMYLGLNGSPSDGTKLVAVTTPVEWHIWHDEVDP
STYRIFVPFTTFNMDLYAQGSAAPGTPITTWYTWKGIHQTWRFELA

SEQ ID NO: 36

TCTCTTCCAGTTTCTACCATGTCGCAAGAAATTGTTCAATCAGGACAAACCTACATCATCACTAACGCCAAATC
CGGCACAGTTGTTGACCTTTCGGGCGAAGACAACAAATCTATTATTGGATTTCCCAAGCATGGAGGAACAAATC
AGAGGTGGACCCTCAACTGGACAGGGAAGAGTTGGACTTTCCGCTCCGTTTCTTCTGAAATGTATCTTGGCCTG
AATGGCTCGCCGTCTGATGGAACAAAACTGGTAGCCGTGACCACCCCTGTTGAGTGGCACATCTGGCACGACGA
AGTTGACCCTTCAACTTATCGTATCTTTGTACCTTTCACCACATTCAACATGGACCTCTACGCCCAAGGTAGTG
CCGCCCCTGGTACGCCTATCACAACTTGGTATACATGGAAGGGCATCCACCAAACGTGGAGGTTTGAACTAGCT
TAGGGTCAGGTTTCGGATGTAATTTGT

SEQ ID NO: 37

TCTCTTCCAGTTTCTACCATGTCGCAAGAAATTGTTCAATCAGGACAAACCTACATCATCACTAACGCCAAATC
CGGCACAGTTGTTGACCTTTCGGGCGAAGACAACAAATCTAGTAAGTCGTTTTTAGTCCCATGTTTTTTTTTGT
CAAAAAAAATTGACTGACATATTTTGTCTCCAGTTATTGGATTTCCCAAGCATGGAGGAACAAATCAGAGGGTA
GGTCTAGAAATGCACCTCGTTAATATTGGTTTTTATTGACATTCATGAACAGTGGACCCTCAACTGGACAGGGA
AGAGTTGGACTTTCCGCTCCGTTTCTTCTGAAATGTATCTTGGCCTGAATGGCTCGCCGTCTGATGGAACAAAA
CTGGTAGCCGTGACCACCCCTGTTGAGTGGCACATCTGGCACGACGAAGTTGACCCTTCAACTTATCGGTGAGT
CCCCTAAATATTACTTGCTTGTGGTTCATACTAATACGTCGTTCGAAGTATCTTTGTACCTTTCACCACATTCA
ACATGGACCTCTACGCCCAGGGTAGTGCCGCCCCTGGTACGCCTATCACAACTTGGTATACATGGAAGGGTATC
CACCAAACGTGGAGGTTTGAACTAGGTAGGGCTTGCGATCTCACCCGGATCCTCCATGAACTAATGTGATCACG
TCGTGTTCTAGCTTAGGTTCAGGTTTCGGATGTAATTTGT

SEQ ID NO: 38

TCGTGCCAGATGCGCCACTCAACAGGGGTGGTCACGGCTACCAGTTTTGTTCCATCAGACGGCGAGCCATTCAG
GCCAAGATACATTTCAGAAGAAACGGAGCGGAAAGTCCAACTCTTCCCTGTCCAGTTGAGGGTCCACCTCTGAT
TTGTTCCTCCATGCTTGGGAAATCCAATAATAGATTTGTTGTCTTCGCCCGAAAGGTCAACAACTGTGCCGGAT
TTGGCGTTAGTGATGATGTAGGTTTGTCCTGATTGAACAATTTCTTGCGACATGGTAGAAACTGGAAGAGAAAA
ACAGTAGGAGCAAAAAAACGAGGCAGAAAGTTCCTCGTGCCGAATTCCTGCAGCCCGGGGGATCCACTAGTTCT
AGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCCTTTAGTA

SEQ ID NO: 39

TCCAACTCTTCCCTGTCCAGTTGAGGGTCCACCTCTGATTTGTTCCTCCATGCTTGGGAAATCCAATAATAGAT
TTGTTGTCTTCGCCCGAAAGGTCAACAACTGTGCCGGATTTGGCGTTAGTGATGATGTAGGTTTGTCCTGATTG
AACAATTTCTTGCGACATGGTAGAAACTGGAAGAGAAAAACAGTAGGAGCAAAAAAACGAGGC

SEQ ID NO: 40

TTTTTTTTTTTTGGGATCCAAAAAGATGTTGCAACATAGGTGGGTATAATGATAACAAGCGGAGTACAATAAAA
GCACAACATGGTCCAAGAAGATTTACACACACAAATTACATCCGAAACCTGAMCCTAAGCTAGTTCAAACCTCC
ACGTTTGGTGGATRCCCTTCCATGTATACCAAGTTGTGATAGGCGTACCAGGGGCGGCACTACCYTGGGCGTAG
AGGTCCATGTTGAATGTGGTGAAAGGTACAAAGATACGATAAGTTGAAGGGTCAACTTCGTCGTGCCAGATGTG
CCACTCAACAGGGGTGGTCACGGCTACCAGTTTTGTTCCATCAGACGGCGAGCCATTCAGGCCAAGATACATTT
CAGAAGAAACGGAGCGGAAAGTCCAACTCTTCCCTGTCCAGTTGAGGGTCCACCTCTGATTTGTTCCTCCATGC
TTGGGAAATCCAATAATAGATTTGTTGTCTTCGCCCGAAAGGTCAACAA

SEQ ID NO: 41

ACAAATTACATCCGAAACCTGACCCTAAGCTAGTTCAAACCTCCACGTTTGGTGGATGCCCTTCCATGTATACC
AAGTTGTGATAGGCGTACCAGGGGCGGCACTACCTTGGGCGTAGAGGTCCATGTTGAATGTGGTGAAAGGTACA
AAGATACGATAAGTTGAAGGGTCAACTTCGTCGTGCCAGATGTGCCACTCAACAGGGGTGGTCACGGCTACCAG
TTTTGTTCCATCAGACGGCGAGCCATTCAGGCCAAGATACATTTCAGAAGAAACGGAGCGGAAAGTCCAACTCT

TCCCTGTCCAGTTGAGGGTCCACCTCTGATTTGTTCCTCCATGCTTGGGAAATCCAATAATAGATTTGTTGTCT
TCGCCCGAAAGGTCAACAACTGTGCCGGATTTGGCGTTAGTGATGATGTAGGTTTGTCCTGATTGAACAATTTC
TTGCGACATGGTAGAAACTGGAAGAGA

SEQ ID NO: 42

ACAAATTACATCCGAAACCTGAACCTAAGCTAGAACACGACGTGATCACATTAGTTCATGGAGGATCCGGGTGA
GATCGCAAGCCCTACCTAGTTCAAACCTCCACGTTTGGTGGATACCCTTCCATGTATACCAAGTTGTGATAGGC
GTACCAGGGGCGGCACTACCCTGGGCGTAGAGGTCCATGTTGAATGTGGTGAAAGGTACAAAGATACTTCGAAC
GACGTATTAGTATGAACCACAAGCAAGTAATATTTAGGGGACTCACCGATAAGTTGAAGGGTCAACTTCGTCGT
GCCAGATGTGCCACTCAACAGGGGTGGTCACGGCTACCAGTTTTGTTCCATCAGACGGCGAGCCATTCAGGCCA
AGATACATTTCAGAAGAAACGGAGCGGAAAGTCCAACTCTTCCCTGTCCAGTTGAGGGTCCACTGTTCATGAAT
GTCAATAAAAACCAATATTAACGAGGTGCATTTCTAGACCTACCCTCTGATTTGTTCCTCCATGCTTGGGAAAT
CCAATAACTGGAGACAAAATATGTCAGTCAATTTTTTTTGACAAAAAAAAACATGGGACTAAAAACGACTTACT
AGATTTGTTGTCTTCGCCCGAAAGGTCAACAACTGTGCCGGATTTGGCGTTAGTGATGATGTAGGTTTGTCCTG
ATTGAACAATTTCTTGCGACATGGTAGAAACTGGAAGAGA

CLAIMS

1. A nematicidal protein which comprises the sequence shown in SEQ ID NO: 1

2. A nematicidal protein having at least 70% identity to a protein as claimed in claim 1.

3. A polynucleotide which encodes a protein as claimed in claim 1 or claim 2.

4. The polynucleotide as claimed in claim 3 which comprises sequence shown in SEQ ID NO: 2.

5. A polynucleotide which is the complement of one which hybridises to a sequence as defined in claim 4 under stringent conditions and wherein said polynucleotide encodes a protein which is nematicidal.

6. A construct comprising the polynucleotide of any one of claims 3 to 5.

7. A construct as claimed in claim 6 wherein the polynucleotide is operably linked to a transcription initiation region and a transcriptional termination region.

8. A construct as claimed in claim 7 which further comprises a selectable marker.

9. A host cell comprising a polynucleotide as claimed in any one of claims 3 to 5, or a construct as claimed in any one of claim 6 to 8.

10. A transgenic plant comprising a polynucleotide as claimed in any one of claims 3 to 5 or a construct as claimed in any one of claims 6 to 8.

11. A method of providing a plant or a plant part with a nematicidal protein comprising:
    a) inserting into the genome of the plant or of plant material a polynucleotide as claimed in any one of claims 3 to 5, or a construct as claimed in any one of claims 6 to 8;
    b) regenerating plants or plant parts therefrom; and

c) selecting those plants or plant parts having said protein.

12. Plants or plant parts obtained according to the method of claim 11.

13. Plants or plant parts as claimed in claim 12 which comprise a further agronomic trait selected from the group consisting of:

a) herbicide resistance;
b) insect resistance;
c) fungus resistance;
d) nematode resistance;
e) altered stress tolerance;
f) altered yield; and
g) altered nutritional content.

14. The use of a polynucleotide as claimed in any one of claims 3 to 5 or a construct as claimed in any one of claims 6 to 8 in a method of producing plants which are resistant and/or tolerant to nematodes.

15. The use of a protein as claimed in claim 1 or claim 2 as an active ingredient in the production of a nematicide.

16. A method of controlling nematodes comprising providing at a locus where said nematodes feed a protein according to claim 1 or claim 3.

17. A composition comprising a nematicidally effective amount of at least one protein as claimed in claim 1 or claim 2 and an agriculturally acceptable carrier and/or a diluent and/or a nematode attractant.

ABSTRACT

IMPROVEMENTS IN OR RELATING TO ORGANIC COMPOUNDS

The present invention relates, *inter alia*, to insecticidal and/or nematicidal proteins. In a particular embodiment the protein comprises the sequence depicted as SEQ ID NO: 1. The invention further provides a polynucleotide encoding the protein of SEQ ID. NO. 1. The invention further provides host cells and plants containing a polynucleotide encoding the protein of SEQ ID NO: 1.

Fig. 1

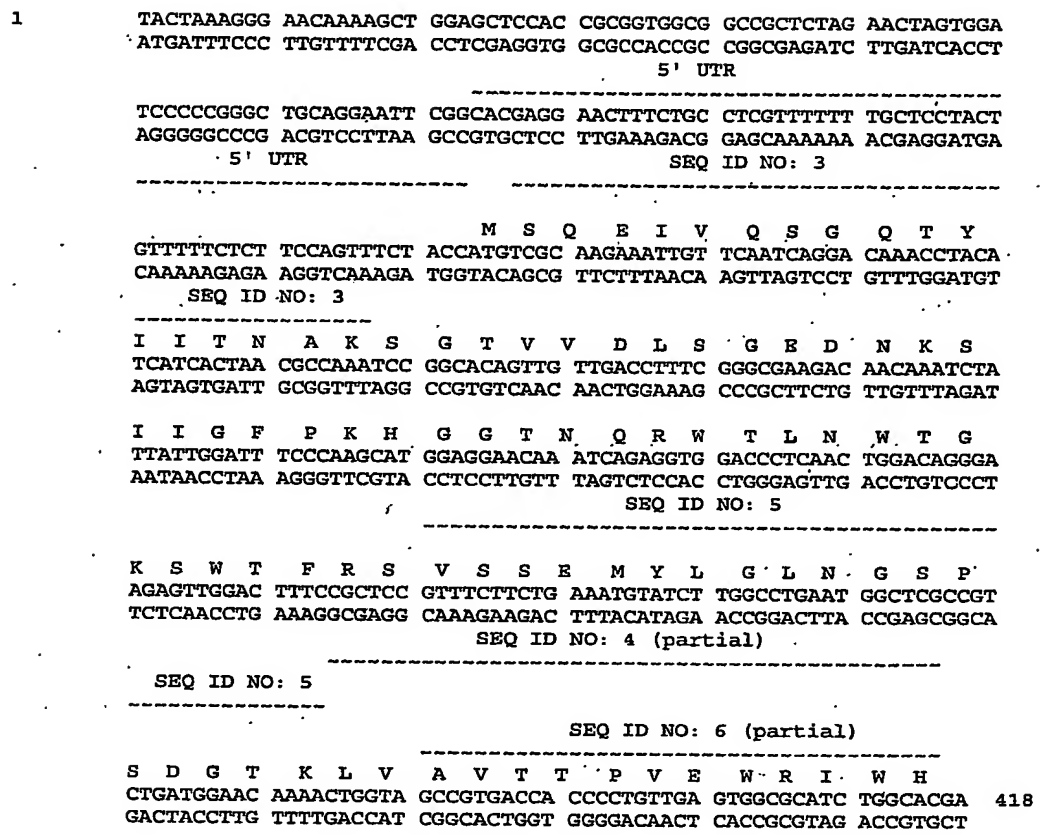

1
TACTAAAGGG AACAAAAGCT GGAGCTCCAC CGCGGTGGCG GCCGCTCTAG AACTAGTGGA
ATGATTTCCC TTGTTTTCGA CCTCGAGGTG GCGCCACCGC CGGCGAGATC TTGATCACCT
5' UTR
TCCCCCGGGC TGCAGGAATT CGGCACGAGG AACTTTCTGC CTCGTTTTTT TGCTCCTACT
AGGGGGCCCG ACGTCCTTAA GCCGTGCTCC TTGAAAGACG GAGCAAAAAA ACGAGGATGA
5' UTR
SEQ ID NO: 3
M S Q E I V Q S G Q T Y
GTTTTTCTCT TCCAGTTTCT ACCATGTCGC AAGAAATTGT TCAATCAGGA CAAACCTACA
CAAAAAGAGA AGGTCAAAGA TGGTACAGCG TTCTTTAACA AGTTAGTCCT GTTTGGATGT
SEQ ID NO: 3
I I T N A K S G T V V D L S G E D N K S
TCATCACTAA CGCCAAATCC GGCACAGTTG TTGACCTTTC GGGCGAAGAC AACAAATCTA
AGTAGTGATT GCGGTTTAGG CCGTGTCAAC AACTGGAAAG CCCGCTTCTG TTGTTTAGAT
I I G F P K H G G T N Q R W T L N W T G
TTATTGGATT TCCCAAGCAT GGAGGAACAA ATCAGAGGTG GACCCTCAAC TGGACAGGGA
AATAACCTAA AGGGTTCGTA CCTCCTTGTT TAGTCTCCAC CTGGGAGTTG ACCTGTCCCT
SEQ ID NO: 5
K S W T F R S V S S E M Y L G L N G S P
AGAGTTGGAC TTTCCGCTCC GTTTCTTCTG AAATGTATCT TGGCCTGAAT GGCTCGCCGT
TCTCAACCTG AAAGGCGAGG CAAAGAAGAC TTTACATAGA ACCGGACTTA CCGAGCGGCA
SEQ ID NO: 4 (partial)
SEQ ID NO: 5
SEQ ID NO: 6 (partial)
S D G T K L V A V T T P V E W R I W H
CTGATGGAAC AAAACTGGTA GCCGTGACCA CCCCTGTTGA GTGGCGCATC TGGCACGA 418
GACTACCTTG TTTTGACCAT CGGCACTGGT GGGGACAACT CACCGCGTAG ACCGTGCT

Fig. 2

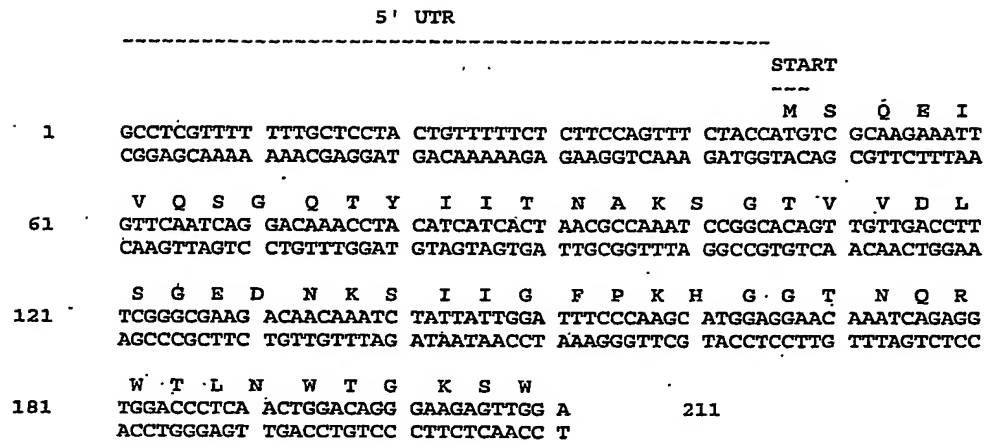

5' UTR
START
M S Q E I
1 GCCTCGTTTT TTTGCTCCTA CTGTTTTTCT CTTCCAGTTT CTACCATGTC GCAAGAAATT
CGGAGCAAAA AAACGAGGAT GACAAAAAGA GAAGGTCAAA GATGGTACAG CGTTCTTTAA
V Q S G Q T Y I I T N A K S G T V V D L
61 GTTCAATCAG GACAAACCTA CATCATCACT AACGCCAAAT CCGGCACAGT TGTTGACCTT
CAAGTTAGTC CTGTTTGGAT GTAGTAGTGA TTGCGGTTTA GGCCGTGTCA ACAACTGGAA
S G E D N K S I I G F P K H G G T N Q R
121 TCGGGCGAAG ACAACAAATC TATTATTGGA TTTCCCAAGC ATGGAGGAAC AAATCAGAGG
AGCCCGCTTC TGTTGTTTAG ATAATAACCT AAAGGGTTCG TACCTCCTTG TTTAGTCTCC
W T L N W T G K S W
181 TGGACCCTCA ACTGGACAGG GAAGAGTTGG A 211
ACCTGGGAGT TGACCTGTCC CTTCTCAACC T

Fig. 3

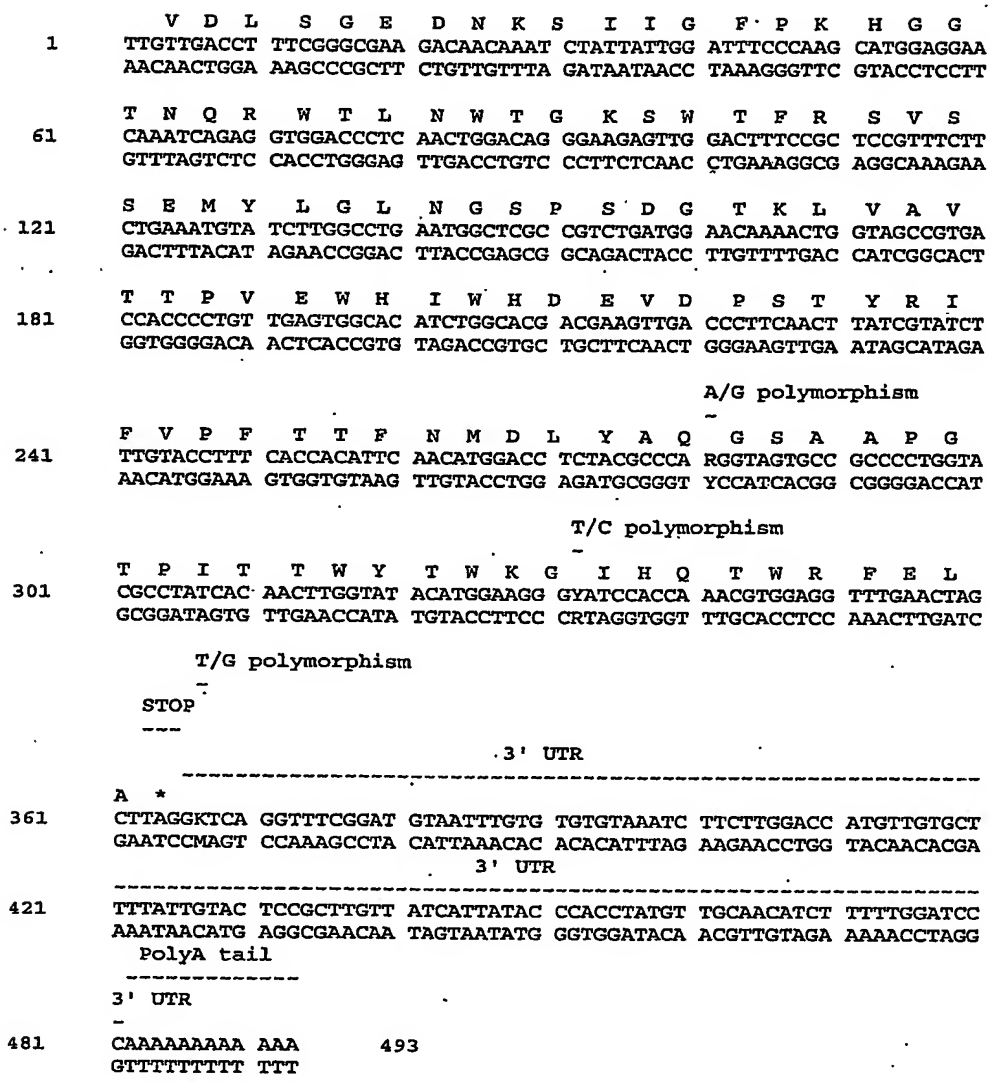

V D L S G E D N K S I I G F P K H G G
1 TTGTTGACCT TTCGGGCGAA GACAACAAAT CTATTATTGG ATTTCCCAAG CATGGAGGAA
AACAACTGGA AAGCCCGCTT CTGTTGTTTA GATAATAACC TAAAGGGTTC GTACCTCCTT
T N Q R W T L N W T G K S W T F R S V S
61 CAAATCAGAG GTGGACCCTC AACTGGACAG GGAAGAGTTG GACTTTCCGC TCCGTTTCTT
GTTTAGTCTC CACCTGGGAG TTGACCTGTC CCTTCTCAAC CTGAAAGGCG AGGCAAAGAA
S E M Y L G L N G S P S D G T K L V A V
121 CTGAAATGTA TCTTGGCCTG AATGGCTCGC CGTCTGATGG AACAAAACTG GTAGCCGTGA
GACTTTACAT AGAACCGGAC TTACCGAGCG GCAGACTACC TTGTTTTGAC CATCGGCACT
T T P V E W H I W H D E V D P S T Y R I
181 CCACCCCTGT TGAGTGGCAC ATCTGGCACG ACGAAGTTGA CCCTTCAACT TATCGTATCT
GGTGGGGACA ACTCACCGTG TAGACCGTGC TGCTTCAACT GGGAAGTTGA ATAGCATAGA
A/G polymorphism
F V P F T T F N M D L Y A Q G S A A P G
241 TTGTACCTTT CACCACATTC AACATGGACC TCTACGCCCA RGGTAGTGCC GCCCCTGGTA
AACATGGAAA GTGGTGTAAG TTGTACCTGG AGATGCGGGT YCCATCACGG CGGGGACCAT
T/C polymorphism
T P I T T W Y T W K G I H Q T W R F E L
301 CGCCTATCAC AACTTGGTAT ACATGGAAGG GYATCCACCA AACGTGGAGG TTTGAACTAG
GCGGATAGTG TTGAACCATA TGTACCTTCC CRTAGGTGGT TTGCACCTCC AAACTTGATC
T/G polymorphism
STOP
3' UTR
A *
361 CTTAGGKTCA GGTTTCGGAT GTAATTTGTG TGTGTAAATC TTCTTGGACC ATGTTGTGCT
GAATCCMAGT CCAAAGCCTA CATTAAACAC ACACATTTAG AAGAACCTGG TACAACACGA
3' UTR
421 TTTATTGTAC TCCGCTTGTT ATCATTATAC CCACCTATGT TGCAACATCT TTTTGGATCC
AAATAACATG AGGCGAACAA TAGTAATATG GGTGGATACA ACGTTGTAGA AAAACCTAGG
PolyA tail
3' UTR
481 CAAAAAAAAA AAA 493
GTTTTTTTTT TTT

Fig. 4

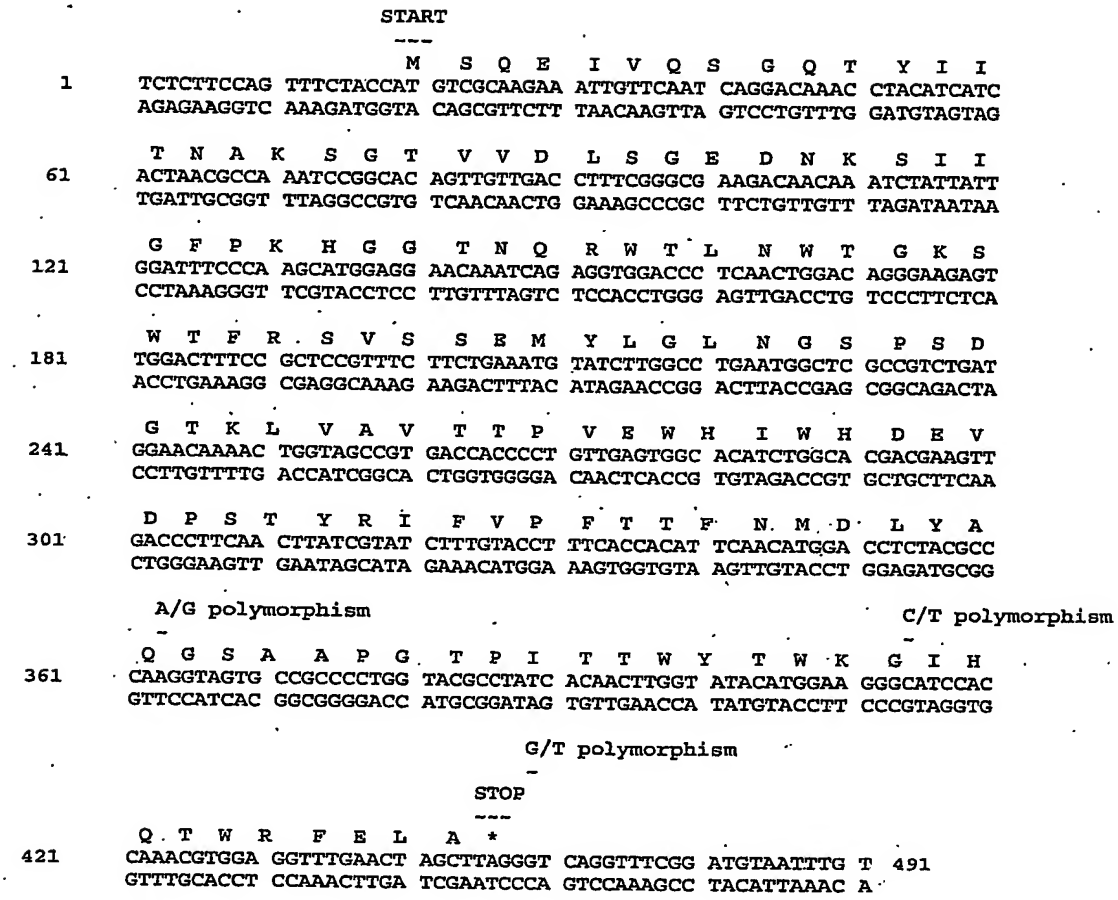

START
M S Q E I V Q S G Q T Y I I
1 TCTCTTCCAG TTTCTACCAT GTCGCAAGAA ATTGTTCAAT CAGGACAAAC CTACATCATC
AGAGAAGGTC AAAGATGGTA CAGCGTTCTT TAACAAGTTA GTCCTGTTTG GATGTAGTAG
T N A K S G T V V D L S G E D N K S I I
61 ACTAACGCCA AATCCGGCAC AGTTGTTGAC CTTTCGGGCG AAGACAACAA ATCTATTATT
TGATTGCGGT TTAGGCCGTG TCAACAACTG GAAAGCCCGC TTCTGTTGTT TAGATAATAA
G F P K H G G T N Q R W T L N W T G K S
121 GGATTTCCCA AGCATGGAGG AACAAATCAG AGGTGGACCC TCAACTGGAC AGGGAAGAGT
CCTAAAGGGT TCGTACCTCC TTGTTTAGTC TCCACCTGGG AGTTGACCTG TCCCTTCTCA
W T F R S V S S E M Y L G L N G S P S D
181 TGGACTTTCC GCTCCGTTTC TTCTGAAATG TATCTTGGCC TGAATGGCTC GCCGTCTGAT
ACCTGAAAGG CGAGGCAAAG AAGACTTTAC ATAGAACCGG ACTTACCGAG CGGCAGACTA
G T K L V A V T T P V E W H I W H D E V
241 GGAACAAAAC TGGTAGCCGT GACCACCCCT GTTGAGTGGC ACATCTGGCA CGACGAAGTT
CCTTGTTTTG ACCATCGGCA CTGGTGGGGA CAACTCACCG TGTAGACCGT GCTGCTTCAA
D P S T Y R I F V P F T T F N M D L Y A
301 GACCCTTCAA CTTATCGTAT CTTTGTACCT TTCACCACAT TCAACATGGA CCTCTACGCC
CTGGGAAGTT GAATAGCATA GAAACATGGA AAGTGGTGTA AGTTGTACCT GGAGATGCGG
A/G polymorphism
C/T polymorphism
Q G S A A P G T P I T T W Y T W K G I H
361 CAAGGTAGTG CCGCCCCTGG TACGCCTATC ACAACTTGGT ATACATGGAA GGGCATCCAC
GTTCCATCAC GGCGGGGACC ATGCGGATAG TGTTGAACCA TATGTACCTT CCCGTAGGTG
G/T polymorphism
STOP
Q T W R F E L A *
421 CAAACGTGGA GGTTTGAACT AGCTTAGGGT CAGGTTTCGG ATGTAATTTG T 491
GTTTGCACCT CCAAACTTGA TCGAATCCCA GTCCAAAGCC TACATTAAAC A

Fig. 5

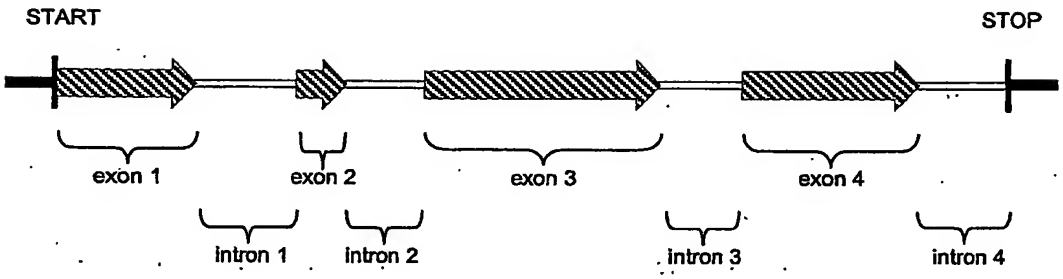

START
STOP
exon 1
exon 2
exon 3
exon 4
intron 1
intron 2
intron 3
intron 4

Fig. 6

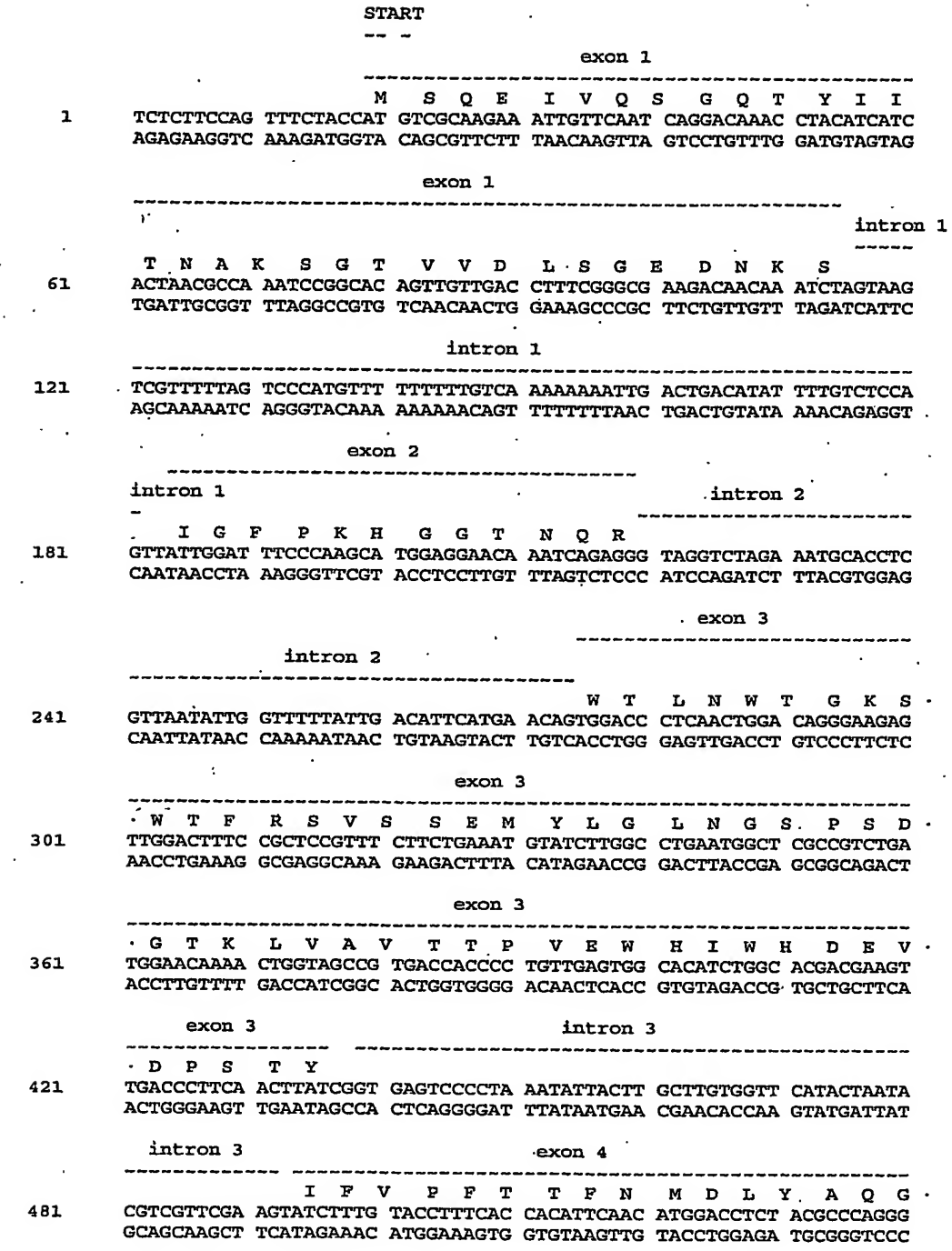
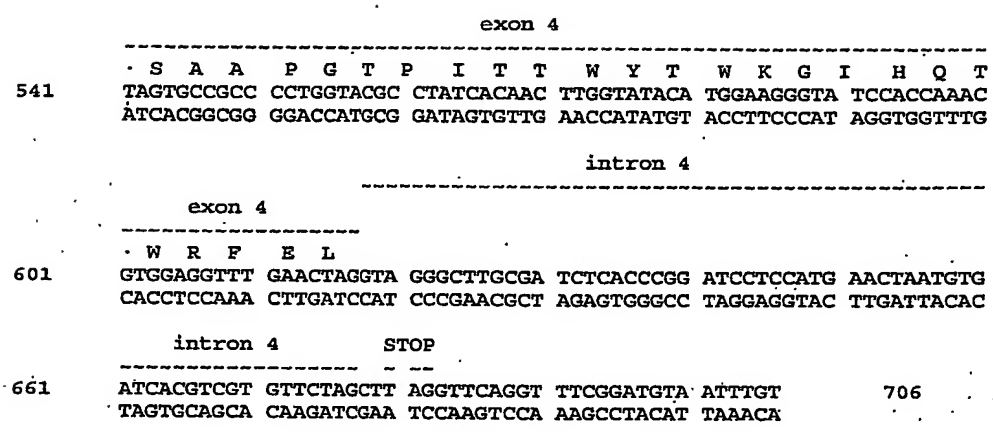

START
exon 1
M S Q E I V Q S G Q T Y I I
1 TCTCTTCCAG TTTCTACCAT GTCGCAAGAA ATTGTTCAAT CAGGACAAAC CTACATCATC
AGAGAAGGTC AAAGATGGTA CAGCGTTCTT TAACAAGTTA GTCCTGTTTG GATGTAGTAG
exon 1
intron 1
T N A K S G T V V D L S G E D N K S
61 ACTAACGCCA AATCCGGCAC AGTTGTTGAC CTTTCGGGCG AAGACAACAA ATCTAGTAAG
TGATTGCGGT TTAGGCCGTG TCAACAACTG GAAAGCCCGC TTCTGTTGTT TAGATCATTC
intron 1
121 TCGTTTTTAG TCCCATGTTT TTTTTTGTCA AAAAAAATTG ACTGACATAT TTTGTCTCCA
AGCAAAAATC AGGGTACAAA AAAAAACAGT TTTTTTTAAC TGACTGTATA AAACAGAGGT
exon 2
intron 1
intron 2
I G F P K H G G T N Q R
181 GTTATTGGAT TTCCCAAGCA TGGAGGAACA AATCAGAGGG TAGGTCTAGA AATGCACCTC
CAATAACCTA AAGGGTTCGT ACCTCCTTGT TTAGTCTCCC ATCCAGATCT TTACGTGGAG
exon 3
intron 2
W T L N W T G K S
241 GTTAATATTG GTTTTTATTG ACATTCATGA ACAGTGGACC CTCAACTGGA CAGGGAAGAG
CAATTATAAC CAAAAATAAC TGTAAGTACT TGTCACCTGG GAGTTGACCT GTCCCTTCTC
exon 3
W T F R S V S S E M Y L G L N G S P S D
301 TTGGACTTTC CGCTCCGTTT CTTCTGAAAT GTATCTTGGC CTGAATGGCT CGCCGTCTGA
AACCTGAAAG GCGAGGCAAA GAAGACTTTA CATAGAACCG GACTTACCGA GCGGCAGACT
exon 3
G T K L V A V T T P V E W H I W H D E V
361 TGGAACAAAA CTGGTAGCCG TGACCACCCC TGTTGAGTGG CACATCTGGC ACGACGAAGT
ACCTTGTTTT GACCATCGGC ACTGGTGGGG ACAACTCACC GTGTAGACCG TGCTGCTTCA
exon 3
intron 3
D P S T Y
421 TGACCCTTCA ACTTATCGGT GAGTCCCCTA AATATTACTT GCTTGTGGTT CATACTAATA
ACTGGGAAGT TGAATAGCCA CTCAGGGGAT TTATAATGAA CGAACACCAA GTATGATTAT
intron 3
exon 4
I F V P F T T F N M D L Y A Q G
481 CGTCGTTCGA AGTATCTTTG TACCTTTCAC CACATTCAAC ATGGACCTCT ACGCCCAGGG
GCAGCAAGCT TCATAGAAAC ATGGAAAGTG GTGTAAGTTG TACCTGGAGA TGCGGGTCCC

Fig. 6 cont...


exon 4
S A A P G T P I T T W Y T W K G I H Q T
541
TAGTGCCGCC CCTGGTACGC CTATCACAAC TTGGTATACA TGGAAGGGTA TCCACCAAAC
ATCACGGCGG GGACCATGCG GATAGTGTTG AACCATATGT ACCTTCCCAT AGGTGGTTTG
intron 4
exon 4
W R F E L
601
GTGGAGGTTT GAACTAGGTA GGGCTTGCGA TCTCACCCGG ATCCTCCATG AACTAATGTG
CACCTCCAAA CTTGATCCAT CCCGAACGCT AGAGTGGGCC TAGGAGGTAC TTGATTACAC
intron 4
STOP
661
ATCACGTCGT GTTCTAGCTT AGGTTCAGGT TTCGGATGTA ATTTGT
706
TAGTGCAGCA CAAGATCGAA TCCAAGTCCA AAGCCTACAT TAAACA

PCT/GB2004/003231